Continuation of a pre-breeding program for improving wheat yield

Johannes Diederick Slabbert

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of Sciences at Stellenbosch University

Supervisor: Willem Botes

Co-supervisor: Marike Visser

March 2020

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2020

Opsomming

Koring is 'n gewas wat al eeue heen wereldwyd verbou word en vorm 'n belangrike deel van basiese voeding, veral in die derde wêreld. Voedselsekerheid is onder groot druk terwyl die bevolking toeneem, en dit is dus belangrik om die hoeveelheid koring wat geproduseer word, te verhoog om aan die vraag te voldoen. Hoër opbrengste en beter gehalte kan bereik word deur 'n verhoging in teelprogramme se effektiwiteit deur navorsing en ontwikkeling van teeltegnieke.

Die doel van hierdie studie was om verhoogde opbrengs verwante eienskappe te identifiseer en dit in die merker bemiddelde herhalende seleksie (MS-MARS) gefasiliteerde voorteelprogram van die Universiteit Stellenbosch Planteteeltlaboratorium (SU-PBL) in te sluit ten einde lyne met hoër opbrengs te teel. Kiemplasma is geïdentifiseer deur middel van literatuur wat die teelprogram kan bevoordeel deur eienskappe te hê wat verband hou met verhoogde opbrengs. Die standaardstel van die SU-PBL se molekulêre merkers is gebruik om ingeligte besluite te neem tydens die seleksie proses. Afstandswaarneming deur 'n ‘Remote Pilot Aircraft System’ (RPAS) is gebruik om gedetailleerde waarnemings van koring uit te voer en dit is vergelyk met tradisionele instrumente. Fenotipering van veld- en na-oes eienskappe is gedoen om te help met die seleksie van lyne met eienskappe met 'n hoë opbrengs.

Lyne met hoë opbrengsverwante eienskappe is geïdentifiseer met behulp van die verskillende tegnieke en is in die MS-MARS-skema gebruik. R-kwadraatwaardes van lineêre regressiemodelle het getoon dat die RPAS-data nie die fenotipiese data in die veld kon voorspel nie, behalwe vir opbrengs. Strooi-plot matrikse het getoon dat daar geen korrelasie was tussen die gegewens wat deur die tradisionele instrumente en die RPAS vasgelê is nie. Die veld- en na-oesdata het aangedui dat 'n naaste buuranalise (NNA) die beste opsie was tydens hierdie studie, aangesien daar veldneigings was en die data is gebruik in die seleksieproses.

Toekomstige studies moet die toevoeging van molekulêre merkers insluit wat ooreenstem met verhoogde opbrengsverwante eienskappe om meer ingeligte besluite tydens seleksie te neem. Die ontwikkeling van kameras en sagteware vir afstandswaarneming sal die instrument beslis bevoordeel. Bykomende vegetatiewe indekse kan ondersoek word en die model kan mettertyd verbeter word deur data by te voeg.

Abstract

Wheat is a crop that has been cultivated around the globe for centuries and forms a substantial portion of the population's diet, particularly in third world countries. Food security is under major stress with the human population increasing, thus it is important to increase the amount of wheat produced to meet the demand. Higher yields along with better quality can be reached by an increased breeding efficiency through research and development of breeding techniques.

The aim of this study was to identify increased yield-related traits and introducing them into the marker-assisted recurrent selection (MS-MARS) facilitated pre-breeding program of the Stellenbosch University Plant Breeding Lab (SU-PBL) to breed wheat lines higher yields. Germplasm was identified through literature that could benefit the breeding program by having traits related to increased yield. The standard set of the SU-PBLs molecular markers were used to make informed decisions during the selection process. Remote sensing by a Remote Pilot Aircraft System (RPAS) was used to perform detailed observations of wheat and was compared to traditional instruments. Field- and post-harvest phenotyping was done to aid in the selection of lines with high-yielding traits.

Lines with high yield-related traits were identified by using the different techniques and were introduced into the MS-MARS scheme. R-squared values of linear regression models displayed that the RPAS data could not predict the phenotypic data in the field, except for yield. Scatter plot matrices shown that there was no correlation between the data captured by the traditional instruments and the RPAS. The field and post-harvest data indicated that a nearest neighbour analysis (NNA) was the best option during this study as there were field trends and the data was used in the selection process.

Future studies should include the addition of molecular markers that correlate with increased yield-related traits to make more informed decisions during selection. The development of cameras and software for remote sensing will definitely benefit the tool. Additional vegetative indices can be explored and the model can be improved over time by the addition of data.

Acknowledgements

I would like to give my gratitude to the following:

• My supervisor, Willem Botes, for all his guidance and inspiration during the course of this study.

• My co-supervisor, Marike Visser, for all her helpfulness and input into this study.

• Stephan van der Westhuizen for his help with the statistical analysis of this project.

• Carlos Poblete-Echeverría from the Department of Viticulture and Oenology at Stellenbosch University for his guidance with the photogrammetry analysis.

• Aletta Ellis and Lezaan Hess, for their technical assistance.

• The students and staff of the SU-PBL.

• The Department of Science and Technology and Grain SA for financial support.

List of abbreviations and SI units

%	Percent
µL	Microlitre
µM	Micromolar
3D	Three dimentional
ABC	ATP binding cassette
ADP	Adenosine diphosphate
$AgNo_3$	Silver nitrate
ANOVA	Analysis of variance
APR	Adult plant resistance
APS	Ammonium persulfate
ATP	Adenosine triphosphate
ATV	All-terrain vehicle
bp	Base pairs
CAPS	Cleaved amplified polymorphic sequence
CCI	Chlorophyll content index
CGIAR	Consultative Group on International Agriculture Research
CIMMYT	The International Maize and Wheat Improvement Center
CIR	Colour-infrared
cm	Centimetre
CMY	Cyan magenta yellow
CO_2	Carbon dioxide
CTAB	N-Cetyl-N, N, N-trimethyl ammonium bromide

CV	Coefficient of variation
DArT	Diversity arrays technology
DH	Doubled haploid
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
DSM	Digital surface model
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
F	Forward primer
FF	Superfine
f. sp.	*Forma specialis*
F1	First generation
F2	Second generation
F4	Fourth generation
g	Gram
GBS	Genotyping-by-sequencing
GNDVI	Green normalised difference in vegetation index
GPC	Grain protein concentration
GPS	Global positioning system
GWAS	Genome wide association studies
g/L	Grams per litre
H^2	Heritability
H_2O	Water
ha	Hectares

HI	Harvest index
hL	Hectolitre
HMW-GS	High molecular weight subunit
HTTP	High throughput phenotyping platform
HYLD	High-yield
i.e.	That is
IPM	Integrated pest management
InDel	Insertion or deletion
IQR	Inter quartile range
IWYP	International Wheat Yield Partnership
kbp	Kilobase pairs
kg	Kilograms
kg/ha	Kilograms per hectare
LAI	Leaf area index
L	Litres
LED	Light emitting diode
LMW-GS	Low molecular weight subunit
Lr	Leaf rust resistance gene
LSD	Least significance difference
Ltd	Limited
LTN	Leaf tip necrosis
M	Molar
MARS	Marker-assisted recurrent selection
MAS	Marker-assisted selection

mg	Milligram
$MgCl_2$	Magnesium chloride
min	Minute
ml	Millilitre
mm	Millimetre
mM	Millimolar
mm^2	Squared millilitre
m^2	Squared meter
MOM	Multilevel thresholding method
Ms	Male sterility
MS-MARS	Male sterility mediated marker assisted recurrent selection
m/s	Meters per second
NaCl	Sodium chloride
NADPH	Nicotinamide adenine dinucleotide phosphate
NaOH	Sodium hydroxide
NaOAc	Sodium acetate
NDVI	Normalized difference in vegetation index
NE	New England
ng	Nanogram
ng/µL	Nanogram per microlitre
NIR	Near-infrared
NJ	Neighbour joining
Nm	Nanometer
NNA	Nearest neighbour analysis

^{o}C	Degrees Celsius
^{o}N	Degrees North
^{o}S	Degrees South
PAGE	Polyacrylamide gel electrophoresis
PAR	Photosynthetically active radiation
PCR	Polymerase chain reaction
pH	Potential of hydrogen
PIC	Polymorphic information content
Pty Ltd	Propriety limited
Px	Pixel
QTL	Quantitative trait loci
R	Reverse primer
R^2	Coefficient of determination
RCBD	Randomized complete block design
RGB	Red, green and blue
Rht	Reduced height
RPAS	Remote pilot aircraft system
rpm	Revolutions per minute
RSA	Republic of South Africa
RVI	Ratio vegetation index
SAGIS	South African Grain Information Service
s	Seconds
SNP	Single nucleotide polymorphism
SPAD	Soil-plant analysis development

Sr	Stem rust resistance gene
SR	Simple ratio
SSD	Single seed decent
SSR	Simple sequence repeat
SU-PBL	Stellenbosch University Plant Breeding Lab
SVI	Spectral vegetation index
T.	*Triticum*
Ta	Annealing temperature
TBE	Tris/Borate/EDTA
TEMED	Tetramethylethylenediamine
TIFF	Tagged image file format
TKW	Thousand kernel weight
Tris-Cl	Tris-chloride
USA	United States of America
UPGMA	Unweighted pair group method with arithmetic mean
USDA	United States Department of Agricuture
UV	Ultra violet
U/µL	Unit per microliter
v	version
VI	Vegetation index
V	Volt
Vrn	Vernalization
v/v	volume per volume
w/v	weight per volume

WES	Welgevallen Experimental Station
WGA	Whole genome assembly
WYCYT	Wheat Yield Collaboration Yield Trial
Yr	Stripe rust resistance gene

Table of Contents

List of figures

List of tables

Chapter 1: Introduction

Triticum aestivum L., common bread wheat, is an economically important crop that is cultivated worldwide, and is an important staple food in many countries including South Africa (Hemdane *et al*., 2015). The crop is cultivated in diverse climatic conditions and is grown on every continent except for Antarctica, even in space (Massa *et al*., 2016). Wheat delivers a primary source of carbohydrates and comprises of proteins, minerals and vitamins. It delivers 20% of the total energy and protein of the food humans consume globally and offers essential carbohydrates, for that reason it is currently being cultivated in 94 third world countries and consumed by more than 4.5 billion people (Braun *et al*., 2010).

The production of wheat has to increase to facilitate the demand due to the growth of the human population. To meet these demands, the wheat should increase by 2.4% every year, while the average increase rate at the moment is only an estimated 1.3% annually. Additionally, there should be a decrease in agricultural inputs and nitrogenous fertilisers to decrease environmental degradation which is caused by the human population's agricultural footprint (Allen *et al*., 2016).

In South Africa, wheat is the second most cultivated crop and is produced in 32 of the 36 crop producing areas. The top producing province in South Africa is the Western Cape (winter rainfall) followed by the Free State (summer rainfall) and Northern Cape (irrigation). These provinces are accountable for 85% of South Africa's entire wheat production (Usda.gov, 2019). The production area of wheat in South Africa has been decreasing for almost 30 years to the present area almost reaching for 500 000 ha planted in the 2018 season. The expected wheat produced in the 2019 season is 1.618 MT, which is 4.54% less than the previous seasons, whilst the expected yield is 3.00 t/ha (SAGIS, 2019). All of the above is possible as a result of having access to the latest research and technology to aid in running an enhanced breeding program to compete against biotic- and abiotic stressors. According to the United States Department of Agriculture (USDA), the net imports of wheat in 2018/2019 stands on 1.8 million tons to meet the demand, but has decreased by 19.05 % compared to the previous year (Usda.gov, 2019). Local plant breeders should work with the goal to

increase grain yield in order to increase production, consequently leading to a decrease in imports.

The aim of this study was to identify yield-related traits and introduce them into the male sterility mediated marker assisted recurrent selection (MS-MARS) facilitated pre-breeding program of the Stellenbosch University Plant Breeding Lab (SU-PBL) to produce and breed wheat lines with high yielding physiological traits by applying cross breeding. In order to reach the aim, the following objectives were recognized:

a) Identifying and obtaining germplasm that can be used to benefit the breeding program.
b) Applying the MS-MARS scheme to contribute to the selection program.
c) The application of high-throughput phenotyping platforms (HTPPs) to perform detailed observations of wheat in an individual and non-invasive manner and to compare them to traditional remote sensing instruments.
d) Perform field and post-harvest phenotyping to select best lines with high-yielding traits.

Chapter 2: Literature review

2.1. Wheat domestication and genome

Wheat and its wild relatives are situated in the *Triticum* genus, which is a member of the ~300 species tribe Triticeae*. Triticum* has six species with the wild taxa existing in the Transcaucasia and Middle Eastern areas. The *Triticum monococcum* L. *(T. monococcum)* and *Triticum urartu (T. urartu)* species provide the AA genome. The *Triticum turgidum* (*T. turgidum*) and *Triticum timopheevii (T. timopheevii)* contributes to bread wheat that consist of the tetraploid AABB and AAGG genome. The last two species are *Triticum aestivum* L. (*T. aestivum)* and *Triticum zhukovskyi (T. zhukovskyi*) that creates the hexaploid AABBDD and AAAAGG genome species. These species are assembled into three divisions that are section Monococcon (2n=2x=14, diploid species); section Dicoccoidea (2n=4x=28 tetraploid species) and section *Triticum* (2n=6x=42, hexaploid species). The *T. zhukovskyi and T. aestivum* occur as cultivated forms and *T. urartu* simply in their wild form. The *T. monococcum*, *T. turgidum* and *T. timopheevii* species have both a cultivated and wild form. The entire *Triticum* genus is native to the 'Fertile Crescent' of the Near East, which consist of the western and northern Iraq, eastern Mediterranean, south-eastern Turkey and its nearby areas northern Iran and Transcaucasus. Introgression from hexaploid wheats have been a crucial mechanism for enhancing and making the gene pool of *Triticum turgidum* wheats more diverse (Matsuoka, 2011) (Table 2.1).

There is substantial variation in morphological and physiological wheat traits, such as grain size, apical dominance, synchronized flowering and growth when the plants are domesticated. These traits are crucial in terms of crop survival and adaption to different environments. *Triticum aestivum* has a complex hexaploid genome that originated through hybridization of the diploid *Aegilops taushcii* and tetraploid *T. turgidum* ssp. Dicoccoides (Allen *et al*., 2016). The genetic diversity in wheat have been decreasing due to domestication, hybridization and pressure caused by selection. The reducing diversity is acknowledged as a restricting factor in the breeding of high yielding varieties, especially with regard to different abiotic and biotic stressors. In order to increase the potential and progress in wheat breeding, it is crucial to evaluate and utilize the genetic diversity of the different germplasms. The assessing of a germplasm

is regularly done on a genotyping level, which aid the exchange of material between different regions and countries in order to establish and mobilize genetic diversity (Reif *et al*., 2005).

An annotated reference sequence that characterize the bread wheats hexaploid genome was sequenced in 2018 that gave access to 107,891 high-confidence genes, this will help breeders and wheat researchers to access information to define changes in the genomes and can help with other technologies such as genome editing (Keeble-Gagnère *et al*., 2018).

2.2. The developmental phases of wheat

The wheat plant experience different phases during development that comprises of germination, tillering, elongation of the stem, boot, heading and ripening. Numerous aspects such as sowing date, temperature, day-length and genotype has an effect on the period of each phase (Acevedo & Silva, 2002).

Various methods can be used to identify the wheat plant phases, the Zadoks scale, for example, is a two-digit system used to classify the developmental phase. The Feekes scale (Figure 2.1) is the most popular scale and describes the eleven phases on a scale from seedling growth to ripening. Seed germination, which can occur within a week from planting, generally start when the seeds absorb sufficient water and when the temperature is between 4°C and 37°C. This results in the presence of seedlings and primary roots as well as the coleoptile that acts as a protective layer of the roots. The coleoptile stops growing and the first leaves starts to develop when the seedling emerges above the soil (Simmons *et al*., 1995).

Table 0.1: Classification of *Triticum*

Section	**Species and subspecies**	**Genome constitution**	**Common names**
Monococcon	*Triticum monococcum* L.	AA	
	subsp. *Aegilopoides*		Wild eirnkorn
	subsp. *Monococcum*		Cultivated einkorn
	Triticum urartu Tumanian ex Gandilyn	AA	
Dicoccoidea	*Triticum turgidum L.*	AABB	
	subsp. *Dicoccoides*		Wild emmer
	subsp. *Dicoccon*		Cultivated emmer
	subsp. *Durum*		Durum or macaroni wheat
	subsp. *Polonicum*		Polish wheat
	subsp. *Turanicum*		Khorassan wheat
	subsp. *Turgidum*		Rivet wheat
	subsp. *Carthlicum*		Persian wheat
	subsp. *Paleocolchicum*		Georgian wheat
	Triticum timopheevii	AAGG	
	subsp. *Armeniacum*		Wild timopheevii
	subsp. *Timopheevii*		Cultivated timopheevii
Triticum	*Triticum aestivum L.*	AABBDD	Common wheat
	subsp. *Aestivum*		Bread wheat
	subsp. *Compactum*		Club wheat
	subsp. *Sphaerococcum*		Indian dwarf wheat
	subsp. *Macha*		
	subsp. *Spelta* (L.)		Spelt
	Triticum zhukovskyi	AAAAGG	

Tillering, the first phase of wheat development, is a crucial phase as the tillers produces heads that are essential for wheat yield. This phase is initiated once the third leaf is developed during the previous seedling phase. The tillers are mainly dependent on the main stem for nutrition and progress from the axils of the key shoot leaves. Although winter wheat has more tillers, not all of the tillers produce grain and will therefore not contribute to a yield increase (Acevedo *et al*., 2002).

Wheat plants continue to develop as temperatures in spring gradually increase and starts to create pseudostems. At this phase the stem has not extended and remains underground. Later during development, the stem will extend and start jointing due to internode elongation. The appearance of a flag leaf, from the whorl, represents the end of the stem extension (Simmons *et al*, 1995). The boot phase is then initiated, and a head will appear and this phase is completed when the awns are visible at the collar of the flag leaf and the sheath opens by the head. The heading phase is identified when there is a complete head visible with two spikelets' located on conflicting sides of a central rachis. Every spikelet contains structures important for reproduction, which are the anthers, ovary, stamen and stigma. Florets that are located on the tillers open a few days after the heading phase and flowers develops both below and above the middle spike. The duration and period of flowering relies on the environmental conditions, where a temperature of 11-13˚C are ideal (Simmons *et al*., 1995, Whingwiri & Kemp, 1980).

Wheat, which is a self-pollinating crop, has a decreased rate of outcrossing since the pollen is shed prior to the opening of the flowers. The pollen that has been shed binds to the stigmatic branches. Fertilization starts when the pollen tube starts to develop after the pollen grain takes up water. The stigma is accessible for 6-13 days, whereas the sustainability of pollen may be less than 30 min in the field (De Vries, 1971). Once the plant has been fertilized the tissues neighbouring the embryo starts to swell which will initiate the growth of seeds. The starch and proteins previously produced as well as products of photosynthesis are relocated to the grain and within ~21 days the grain will have their highest fresh weight (Wan *et al*., 2008).

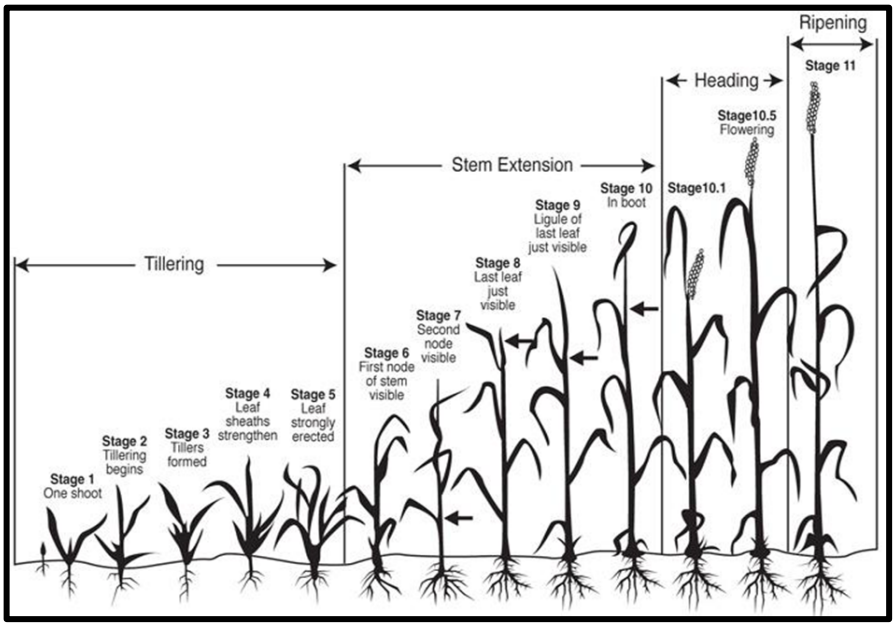


Figure 0.1: Developmental stages of wheat plants (Marsalis & Goldberg, 2016).

2.3. Factors restricting wheat cultivation

Farmers' worldwide experience challenges regarding the cultivation of wheat due to numerous factors. Biotic and abiotic stressors have a considerable influence on crop yields and can have a direct or indirect effect on decreased yield and quality of the crops while plant diseases cause an annual loss of 10-16% of the global harvest (Frenkel *et al*., 2017).

2.3.1. Biotic stressors

Biotic stresses that have an effect on wheat crops include insects or pathogenic infections by plant viruses, bacteria and fungi. The Russian wheat aphid, mites and stalk bores are insects that generally infest wheat in South Africa, while rye grass is the most common weed that has a negative effect on wheat yield in the Western Cape. One of the leading biotic stresses that have an impact on local wheat production is the diseases triggered by the *Puccinia* fungi (Duveiller *et al*., 2012). *Fusarium graminearum* is a pathogen that can cause the gathering of mycotoxins (metabolites with low molecular weight) in the grain that can harmful to the consumer (Desjardins & Hohn, 1997).

2.3.2. Abiotic stressors

Abiotic stresses are identified as non-living factors that have a negative influence on the production of wheat (Atkinson & Urwin, 2012). Examples of abiotic stressors comprise of water stress, severe temperatures, deviation of inorganic solubles from the optimal concentration, irregular photon irradiance and ground acidity and salinity. Numerous abiotic stresses might be present under certain circumstances at the same time and can influence the capability of the wheat to dodge the stress (Abhinandan *et al*., 2018).

2.3.3. Tolerance

A plant's tolerance can be described as its capability to survive, grow, develop and reproduce in unsuitable conditions. The tolerance of certain plant species can be complex due to the fact that different plant species each have different ways to tolerate certain biotic and abiotic stresses. The difference in tolerance capability is due to the different interactions between the plant's molecular, physiological and biochemical structures and stressors. In the case of wheat, the cultivars can be enhanced in order to create drought resistance by selecting for alleles from the wild relatives of the wheat (Oyiga *et al*., 2017).

2.4. Approaches to increase wheat yield in South Africa

2.4.1. Breeding programs

Plant breeders previously used to identify resistant genes visually based on observing the phenotypical characteristics, however later realised that supplementary factors had an influence on the plants and that they cannot be assessed only on a phenotypic level. These supplementary factors were identified as having low repeatability and the genes of interest were being disguised by other additional genes. Biotechnology has a major influence on the productivity of plant breeding due to the fact that this field of work makes it possible for farmers and breeders to have more knowledge on the trait of interest not just phenotypically, but also on a molecular level (Breseghello & Coelho, 2013).

During early cultivation and domestication, selection had a crucial role in the wheat improvement leading to the current high yielding cultivars that possess a hexaploid genome and is seen as the origin of plant breeding by wheat researchers. The productivity of wheat changed a lot since the first selections up until now due to research and breeding. The green revolution was one of the milestones that changed the wheat plant the most as it increased the production in developing countries (Khush, 2001).

2.4.2. The International Maize and Wheat Improvement Center

The International Maize and Wheat Improvement Centre (CIMMYT) is the frontrunner of the Global Wheat Program of the Consultative Group on International Agriculture Research (CGIAR) (CGIAR, 2020). One of their main goals is to improve the efficiency of wheat cropping systems in order to assist in decreasing poverty in developing countries. CIMMYT's wheat program also aims to develop and improve new wheat germplasm which are open for use by national and international partners in advancing their own germplasm, and in some cases to be released as varieties (Cimmyt.org, 2018).

New breeding lines are annually assessed. The best lines are selected and bred for different environments before being distributed to the different national partners. The

germplasm that CIMMYT produce is used around the world especially in developing countries (Guzman *et al*., 2016).

2.4.3. International Wheat Yield Partnership

The International Wheat Yield Partnership (IWYP) is a worldwide attempt that ensures that there is durable funding from private and public research organizations. These funds are used for coordinating research to aid in improvement of scientific research. This partnership was established in 2012 and has been growing rapidly ever since with the main aim is to increase genetic potential of wheat by 50% by 2025. The partnership focuses on six main research sections; 1) the discovery of genetic variation responsible for changes in the fixation of carbon and segregate amongst wheat lines; 2) coupling genes from wheat and additional species through genetic modification to increase the biomass production by enhancing carbon fixation; 3) improving wheat growth and development to enhance grain yield; 4) to produce elite wheat lines that can be included into other breeding programs; 5) using the available information on wheat relatives and other species to improve wheat lines; 6) using new technology and techniques that can improve wheat breeding programs (Iwyp.org, 2020).

2.4.4. Pre-breeding programs

Breeding lines have unfavourable genetic history or low agronomic traits and it will be a drawback to introduce unwanted genes into a breeding program. The genetic variability in wild relatives is frequently held up by linkage drag or the incompatibility between wild relatives or breeding lines. These problems can be resolved through the use of pre-breeding programs which can support the use of genetic variability in the incompatible wild type germplasm. This program will assist the transfer of desirable traits from a contributing germplasm to an uncomplimentary breeding population. The material can be used to develop new varieties with sufficient genetic variability to advance breeding programs and produce new superior lines (Gorjanc *et al*., 2016).

2.4.4.1. Breeding for agronomical characteristics

Plant breeders need to continuously react to changes in different environmental circumstances, agricultural practises and consumer preferences. The different practises should be researched when breeding new cultivars with desired characteristics.

2.4.4.2. Breeding for wheat quality end-use

Wheat products are important sources of proteins and energy. Wheat can be classified into two groups, soft wheat and hard wheat. Soft wheat has a soft endosperm and is mainly used for crackers, pastries, cakes and flat bread. Hard wheat, has a harder endosperm than soft wheat and is mainly used for leavened and unleavened breads ("pieta", "tortillas" and "chapattis") (Atwell, 2001).

During the Green Revolution the production of wheat increased, which led to chief import countries having a production shift becoming net export countries such as Mexico and India. Currently, these countries have the objective to export wheat as well as to compete in the global grain market constructing quality products. Since the beginning of the Green Revolution, the main goal of wheat breeders is to produce more stable wheat with higher yields (Trethowan *et al*., 2007). The use of high-yielding semi-dwarf wheat cultivars led to increases in grain yield but decreased the protein concentration. This was mostly a result of negative correlation amongst the two traits, which originated from a dilution effect of protein between a large amount of carbohydrates (Ortiz-Monasterio *et al*., 1997).

The proteins in wheat grains essentially consist of glutenin, albumin, gliadin and globulin. Gliadin and glutenin, together with energy in the form of calories and water forms gluten, which is the most important factor in the quality of flour and bread making. Albumin and globulin do not play a crucial role in the quality of bread making, however it is important for the metabolism of the plant (Gao *et al*., 2009).

2.4.4.3. Breeding for resistance

A crucial objective for the majority of plant breeders is the ability to breed for resistance against biotic or abiotic stressors and to ensure that the plant can survive, grow and reproduce (Garcia *et al*., 2013). The initial step is to identify a resistant germplasm that can overcome an essential stressor, followed by introducing this into a population that has agronomical desirable traits. The resistant line should then be cultivated in different locations and must be robust. There are a variety of different breeding methods for the incorporation of resistant genes currently being applied, such as pedigree selection, bulk selection, single seed decent, backcrossing and recurrent selection (Acquaah, 2015).

2.5. The use of biotechnology in breeding programs

2.5.1. Genetic markers

2.5.1.1. Different genetic markers used in wheat breeding

Plant pre-breeding programs have included modern biotechnology tools that increased the productivity and speeded up procedures. The use of identified molecular markers to locate the position of genes that are desired in a breeding crop is known as marker-assisted selection (MAS). These markers are linked to regions on the chromosome better known as quantitative trait loci (QTL) or the target gene in a close or direct manner (Chhetri *et al*., 2017). As soon as the traits are mapped, a closely linked marker can be used to screen large numbers of samples to detect progeny that includes some of the desired features. MAS is a technique to select genes that have desired traits based on genotypes, instead of phenotypes. This technique can be used to increase the rate of development of different wheat cultivars and to differentiate between different genes for similar features (Rana *et al*., 2011). When a trait has a lower heritability, the use of MAS can be valuable. The traits that indicate little heritability will be seen less in a population when the detection of QTL gets challenging, or when the reliability of the markers are influenced by the interaction of the environment on the QTL (Bruno *et al*., 2017). MAS is commonly used in commercial breeding programs in order to increase and improve grain by predicting the breeding value of a singular germplasm by selecting and backcrossing desired alleles to create an elite germplasm.

MAS also get used in breeding programs especially in cases where analysis has to be done on a large amount of seed. The use of MAS is argued to be a promising method to use in order to make premature selections based on the plant's DNA based markers and protein profiles (Hospital, 2009).

MAS is furthermore used in the male sterility marker-assisted recurrent selection (MS-MARS) scheme which is a repeated selection scheme due to being inexpensive and effective. This scheme was proposed by Marais *et al*. (2000) and comprised of breeding cycles where the females and males were treated in a different manner. They introduced the dominant male sterility gene, *Ms3* into spring wheat "Inia 66" that was derived from the male sterile winter wheat accession KS8/UP9. The introduction of this gene leads to the recurrent selection base population having a 50% male sterility for each one of the recurrent cycles and insures that the female plants are easily obtained. The main objective of this scheme is to increase the frequency of the desired genes within the base population (Marais & Botes, 2009) (Figure 2.2).

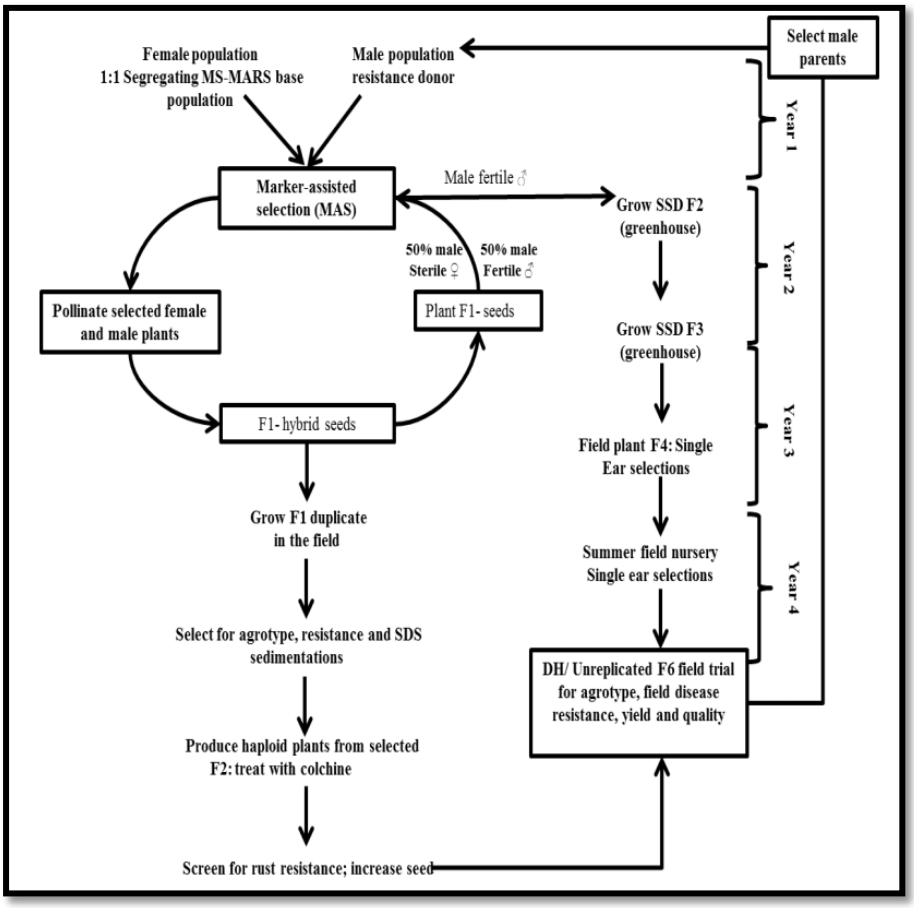


Figure 0.2: Male sterility marker-assisted recurrent selection scheme (Marais & Botes, 2009).

The markers can be used to target novel traits from different germplasm and to detect recombinants. Single nucleotide polymorphism (SNP) and other types of molecular markers are used for selection to improve the specific crop. SNPs are used for high throughput analysis; however very little SNP information is available for wheat. The technology has progressed at a speedy pace, while some of the molecular markers have been employed for years (He *et al*., 2014).

QTLs are the areas within genomes that hold genes that are related to a specific quantitative trait and it cannot be recognized by only evaluating the phenotype. Numerous QTLs have been identified that can be used to enhance crop performance in environments with different challenges such as water stress (Oyiga *et al*., 2017).

Microsatellites or simple sequence repeats (SSRs) are tandem repeats that are spread throughout the genome and consist of repetitive monomer sequences (1 to 6 base pairs in length) that can be amplified by the utilization of polymerase chain reaction (PCR). In earlier studies, the application of microsatellite markers has set a worthy platform for the application of QTL mapping and consecutively MAS in breeding programs throughput (Dreisigacker, 2012).

Diversity arrays technology (DArT) is applied to identify changes in indels and single bases in the genome without using a reference sequence. It is cost effective, high-throughput method, which can be used to map the whole genome of a crop (Dreisigacker, 2012).

Genotyping-by-sequencing (GBS) is a highly multiplex system that can effortlessly construct reduced representation libraries that is needed for the Illumina next-generation sequencing (NGS) platform. This platform produces a great amount of SNPs that can be used in genotyping. The benefits of this platform include that it is less expensive, demands less effort and PCR and has no limit on reference sequences. One of the most significant applications of GBS is the use of it in plant breeding since it allows plant breeders to apply genome wide association studies (GWAS), linkage analysis, genomic selection, diversity studies and the discovery of molecular markers. Prior knowledge of the genome is not needed as the method proved to be robust across multiple species and the detection of SNPs and genotyping are done simultaneously (He *et al*., 2014).

2.5.1.2. Genetic markers as a selection tool

Genetic markers are short DNA sequences that can be used to characterize genetic variation between individual organisms or species, while they perform as a 'flag' to identify target genes. It identifies variations in the genomic loci that can be due to mutations and modifications. The benefit of using these markers is that this is an inexpensive process that is not easily influenced by the environmental factors (Spindel *et al*., 2015). These markers are then used to aid as a tool in selection by association or linkage mapping analysis, however it must initially be approved by breeding populations. The success of the markers in a breeding program is determined by its ability to be integrated into the population. Consequently, the utilization of MAS in wheat breeding programs is used worldwide, depending on the program goals and the available information and resources (Dreisigacker, 2012; Desta & Ortiz, 2014).

Wheat around the globe is constantly subjected to fungal diseases caused by three different wheat rusts, known as leaf rust (*Puccinia triticina*), stripe rust (*P. striiformis f.sp. tritici*) and stem rust (*P. graminis f.sp. tritici*) (Sucher *et al*., 2016). The most environmentally friendly and effective way to control and regulate fungal diseases, while improving yield, is grounded on new cultivars containing resistant genes (Chauhan *et al*., 2015). Several resistant genes have been discovered that were, however, pathotype-specific and are easier to be overcome by pathogens that evolve into novel virulent pathogens. Resistance breeding against rust transformed by focusing more on attaining durable resistance which is non-pathogen specific and will extent the effects of the resistance (Ellis *et al*., 2014).

The collection of rust resistance genes can be split up in two categories namely seedling resistance genes, which is mainly expressed throughout the seedling phase of the plants lifecycle, or the adult plant resistance (APR) genes, which is expressed during the mature phase of the plant's lifecycle. The seedling resistant genes are mainly pathotype specific, while the APR genes are mainly non-pathotype specific resistant genes (Ellis *et al.,* 2014).

Since 1966, the *Lr34* gene, located on chromosome 7DS, was identified and is known for being one of the several slow rusting genes that will contribute to durable resistance. When wheat plants experience necrosis, it indicates the presence of *Lr34*

since the gene is associated with the leaf tip necrosis (LTN) locus. This gene has been used effectively for more than a century and also delivers resistance to barley yellow dwarf virus and powdery mildew (Keller *et al*., 2013). It is known that the *Lr34* gene encodes for an ATP-binding cassette (ABC) transporter, which is connected to transport of the active transmembrane of different biochemical compounds. Since the 1920's the *Sr2* gene is an important gene in breeding programs for the resistance against stem rust and is also a non-pathotype specific APR. The gene is located on the 3B chromosome of the wheat genome and was initially transferred from wild emmer wheat. *Puccinia* pathotypes has not overcome either of the *Lr34* or *Sr2* genes and is, therefore, used globally in wheat breeding programs in order to build rust resistance in novel cultivars (Lagudah, 2011).

Molecular markers that are associated with several resistance genes are used in the screening process of new wheat cultivars to identify possible resistance genes. It is important to validate the molecular markers that are being used in wheat breeding programs occasionally by presenting inoculation tests on wheat plants that are known to have the genes of interest. This confirms that the molecular markers are segregating with the resistance genes and that recombination did not detached the gene from the molecular marker. This aids in preserving the resistance genes in the germplasm (Krattinger *et al*., 2009).

An ideal breeding program contains wheat lines with a maximum number of resistance genes along with other favourable agronomical traits. Bread is one of South Africa's staple foods and is also the main means of wheat consumption. One of the most important factors that breeders select for is the baking quality of the flour produced from the wheat plants. The gluten proteins, found in the endosperm of the grain, primarily regulates the baking quality, since it has an influence on the water retention capabilities and the visco-elasticity of the dough. The baking quality of 'SST027' wheat cultivar (Sensako, South Africa) forms the standard to which novel cultivars are compared. Gluten is formed by both components; gliandin and glutenin. Glutenin has either a low molecular weight subunit (LMW-GS) or high molecular weight subunits (HMW-GS). The standard marker set used by SU-PBL has molecular markers for both LMW-GS and HMW-GS because both these sub-units regulates the protein and baking quality of the flour (Xu *et al*., 2008).

Both of the reduced height genes *Rht-B1b* and *Rht-D1b* (previously known as *Rht1* and *Rht2*) are included in the standard panel of SU-PBL and represent different semi-dwarfing phenotypes in wheat. The genes lead to wheat genotypes with stronger and shorter stems that can reduce wheat with heavy grains to lodge that leads to cultivars with higher yield potential. These genes were introduced into American wheat germplasm from “Norin10", a Japanese cultivar. Molecular makers for the alleles for these genes were designed and are currently in more than 70% of the world’s cultivars that are planted commercially (Langridge, 2014).

Four yield-related molecular markers were added to the panel of markers at the SU-PBL used for studies associated with yield (Rhoda, 2018). These markers represented the *TaGS5-3A*, *TaGW2-6B*, *TaGS-D1* and *Ppd-D1* genes. The *TaGS5-3A* (Ma *et al*., 2015) and *TaGW2-6B* (Qin *et al*., 2014) genes is a positive regulator for grain width and weight and their presents can be linked to an increase yield. The *TaGS-D1* gene is correlated with having in increased grain length and weight (Zhang *et al*. 2014) and the *Ppd-D1* gene to an earlier heading date which let the wheat be ready for harvest at an earlier stage before the summer days are too warm (Guo *et al*., 2009).

2.6. Production of doubled haploids

Conventional breeding programs are thought to be very time-consuming since they require numerous rounds of selection, inbreeding and crossing. It needs to undergo the homozygotisation process, resulting in the registration of a novel cultivar taking several years. The double haploid method can be used to develop homozygous lines from former heterozygous lines in a shorter time (Ślusarkiewicz-Jarzina *et al*., 2017). Doubled haploids (DH) of crops can be attained from hybrid plants and have complete homozygosity and is used to increase the cultivar development in genetic manipulation, gene mapping, crop improvement and the genome. There are two key stages namely haploid induction and chromosome doubling. An effective protocol for inducing haploids are crucial for creating mapping populations and the breeding of commercial cultivars (Niu *et al*., 2014).

2.6.1. The use of doubled haploid technology in breeding programs

The classification of a DH is a genotype that is created when haploid (n) cells effectively experience chromosome doubling, either artificially induced or randomly. This process causes a significant reduction in the time of the breeding cycle through the quick production of homozygous lines (2-3 generations), rather than conventional inbred line production (6-8 generations) that create lines with approximately 99% homozygosity (Prasanna *et al*., 2012). This is an efficient way to develop new wheat cultivars since it is self-pollinating species. The introduction of the traits of interest is achieved faster and more efficiently when compared to conventional breeding due to the absence of dominance variance, higher genetic variance and decreased environmental influences. Haploids can appear naturally in higher plants but are artificially produced for creating new cultivars and can be done *in vitro* or *in vivo* through the use of androgenesis and gynogenesis (Niu *et al*., 2014).

2.6.2. Different ways to produce doubled haploids

2.6.2.1. Wide crossing

Wide crossing or hybridization in wheat is the crossing of wheat with other crops to produce new hybridized seeds. It has been utilized by crossing wheat with maize, barley, sorghum, pearl millet, and teosinte and can successfully produce diploid wheat plants (Niroula & Bimb, 2009). The pollination of wheat by maize commonly produces embryos. It was cytologically proved that maize pollen usually germinates wheat which then can develop into the embryo sac where the wheat's ovum is fertilized by the sperm of the maize. This fertilization step creates a hybrid zygote that comprise of 10 maize and 21 wheat chromosomes. This will result in the failed movement of maize chromosomes towards the spindle poles due to the karotypically unbalanced hybrid zygote. Furthermore, it will result in the elimination of maize chromosomes after a number of cell divisions that will produce a haploid embryo with 21 wheat chromosomes plants (Niroula & Bimb, 2009).

2.6.2.2. Gynogenesis

In normal conditions the egg cell needs to be fertilized to develop into an embryo. Nevertheless, the cell entails the genetic information needed to start embryogenesis without fertilization (parthogenesis). In selected apomictic plants; the female egg divides without being fertilized by the male sperm. In a few other plants the egg cell can be produced and divided by external treatments (Bhojwani & Dantu, 2013). In the case of gynogenesis, the first few cell divisions are almost identical as in zygotic embryogenesis. However, the egg cell is involved in cell divisions that are organized and forms proembryos as well as well-differentiated embryos. This method is usually used in the case of producing haploids on a smaller scale since the amount of megaspores of wheat are low and female haploid cell isolation is challenging (Niu *et al*., 2014).

2.6.2.3. Androgenesis

Androgenesis is a method that is used to create DHs and was first used more than thirty years ago on wheat and barley. Despite all the improvement in the methodology that is used to produce more divisions of microspores in crops, there are factors that inhibits the success of this technique in breeding programs (Echávarri & Cistué, 2015). Some of the factors during production of haploids when using anther or microspore culture include the genotype response, high albinism rate and the prolonged periods during regeneration which will lead to gametoclonal variation and mixed-ploidy crops (Niu *et al*., 2014).

Androgenesis comprises of three important steps: pre-treatment, induction and regeneration. During pre-treatment the microspores are stressed under different chemical, nutritional and environmental conditions to introduce embryogenic development. During induction the embryos should be established by the embryogenic microspores. The regeneration step is the germination of the consequential embryoids that originated from the microspores (Rout *et al.,* 2016).

2.6.2.4. Physiological trait breeding

Physiological trait breeding concentrates on the advancement of crop and will help them to adapt to different abiotic stresses with the aim to increase the potential of wheat yield. Physiological traits can be used in selection to breed improved lines in order to breed wheat to meet the increased worldwide demand (Garcia *et al*., 2013). The majority of breeding programs concentrates on selecting main resistance genes and the introgression of dwarfing or novel genes into a new germplasm. The major areas that physiological research are applied in breeding programs are (i) physiological characterisation of possible parents that can be used in crossing to gather genes that can adapt to stress; (ii) reintroducing progeny from early generations that holds a desired trait to increase the frequency of the gene and also to exclude lines that holds undesired traits (iii) genotypes that occupies favourable physiological traits can possibly be acknowledged by the use of genetic resources (Reynolds *et al*., 2012).

2.7. Breeding for agronomical characteristics

Plant breeders need to continuously react to changes in different environmental conditions, agricultural practises and preferences made by the consumers. The different practises should be researched when breeding new cultivars with desired characteristics (Ceccarelli, 2015).

2.7.1. Water and temperature relations traits

The majority of physiological processes are dependent on the water conditions and can regulate the growth and development of the crops. A cultivars potential to excel can be measured by the plants water status. There are various methods to screen for physiological traits that represent the water status of the plants, for example: chlorophyll retention, canopy temperature, carbon isotope discrimination, stomatal conductance, osmotic adjustment, roots and relative water content (Cossani *et al.,* 2012).

2.7.2. Chlorophyll retention

Photosynthesis, a process which occurs in the plant's chloroplast, is the second most important process that synthesize ATP (Adenosine 5'-triphosphate) which is the main molecule for relocation and storage of energy in cells. The pathway's main end products are two carbohydrates; a disaccharide sucrose and leaf starch. The leaf starch is produced and stored in the chloroplast, where the sucrose is produced in the cytosol and sent to different parts of the plant. Photosynthesis also produces oxygen which is then released through the stomata of the plant (Lodish *et al*., 2000).

The plant's photosynthesis pathway consists of four stages that takes place in the chloroplast which are: (1) light absorption, (2) the reduction of $NADP^+$ to NADPH through electron transport, (3) production of ATP and the (4) formation of carbohydrates from CO_2. The light absorption step is done by the chlorophylls that are attached to proteins in the thylakoid membranes. Oxygen is formed through the removal of electrons from water (unwilling donor) and is done by way of the energy of the absorbed light. During the electron transport step in the thylakoid membrane, the electrons are transferred from the quinone primary electron acceptor up until they reduce $NADP^+$ (electron acceptor) to NADPH. Throughout the production of ATP the protons passes down their concentration gradient when it goes through the F0F1 complex from the thylakoid lumen to the stroma, which synthesizes ATP from ADP and P_i. During the carbon fixation step the ATP^{4-} and NADPH produced in the previous steps delivers the energy and electrons necessary to synthesize CO_2 and H_2O (Lodish *et al*., 2000).

Photosystems, located in the thylakoid membrane, are multiprotein complexes responsible for the absorption of light energy and the transformation to chemical energy. They have two related mechanisms, an antenna comprising of light-absorbing pigments and a reaction centre consisting of proteins and a chlorophyll molecule. Chlorophyll a is the main pigment in organism's that experience photosynthesis and is existent in both antennas and reaction centres. Antennas contain three light-absorbing pigments that include chlorophyll a, chlorophyll b and carotenoids that absorb light at different wavelengths (Lodish *et al*., 2000).

The green region presented by a plant is a good indicator of its photosynthetic capacity, which is the maximum speed at which the plant is able to fix carbon. Chlorophyll retention is considered as a crucial indicator of the adaption to stress (Lopes & Reynolds, 2012). Senescence is a genetically programmed process that is influenced by the environment that leads to the destruction of chlorophyll and utilization of nutrients to newer or reproductive parts of the crops. Chlorophyll retention is a form of delayed senescence and the phenotype has confirmed usefulness to aid in improving yields under biotic and abiotic stress. The expression of stay-green has a major advantage in yield under post-anthesis drought compared to hybrids that does not have this trait. The delay of leaf senescence is linked to higher grain yield or protein concentration; however it relies on the environment (Hörtensteiner, 2009).

The stay-green phenotype is usually allocated by visual annotations as the plant has a greener appearance but can also be done by identifying retention of the green leaf area and how fast the plant is losing chlorophyll with SPAD (soil-plant analysis development) (Lopes & Reynolds, 2012). Normalized difference vegetation index (NDVI), a remote sensing measurement used to indicate if the target holds live green vegetation, and additional indices from multispectral sensors can be used to estimate chlorophyll content, nitrogen-status, biomass, ground-cover and yield in wheat. The NDVI also provides a continuous measurement of chlorophyll retention and can replace discrete scores achieved by subjective visual observations (Lopes & Reynolds, 2012). Stay-green can be correlated with yield under heat stress and a combination of drought and heat by the use of NDVI stress (Montazeaud *et al*., 2016).

There is an inverse genetic relationship when concurrently breeding for grain protein concentration (GPC) and grain yield traits in wheat. The association is a result of interactions of nitrogen and carbon metabolism that individually affect the starch or amino acid synthesis. Considering the involvement of nutrients stored prior to anthesis to the final yield (57% in wheat), it is clear that a major part of the final grain yield originates from post-anthesis translocation and accumulation of nutrients, mostly carbon and nitrogen. Both GPC and grain yield are strongly influenced by the length of the post-anthesis period and conservation of green leaf as an indication of photosynthesis. GPC and grain yield are, therefore, determined during leaf senescence post-anthesis. Studies have identified wheat which expressed early senescence and chlorophyll retention which poses a valued resource to aid in

improving GPC and grain yield of winter wheat cultivars. A positive correlation has been made between yield and chlorophyll retention phenotypes. Further large-scale phenotyping studies under field conditions of premature senescence and chlorophyll retention will, therefore, improve our knowledge of the effect of senescence on GPC and grain yield (Kipp *et al*., 2014).

During the start of senescence there is a reduction in the plants chlorophyll content, and this is used to distinguish between early senescence and chlorophyll retention phenotypes. The disadvantage of using this measurement is that the plant does not experience senescence in a uniform manner, where the older leaves undergo senescence before the younger leaves. There are, however, alternative indicators in the flag leaf area, resulting from nitrogen uptake and carbon remobilization, which can be used to identify the start of senescence (Havé *et al*., 2015)

2.7.3. Canopy temperature

Canopy temperature is the amount of energy being released and is expressed as the temperature of the crop that can deliver information on the plants metabolic function and water usage and amount (Figure 2.3). Infrared thermometers can be used to measure plant temperature, which is a valuable qualitative index to differences in plant water usage (Jackson *et al*., 1981).

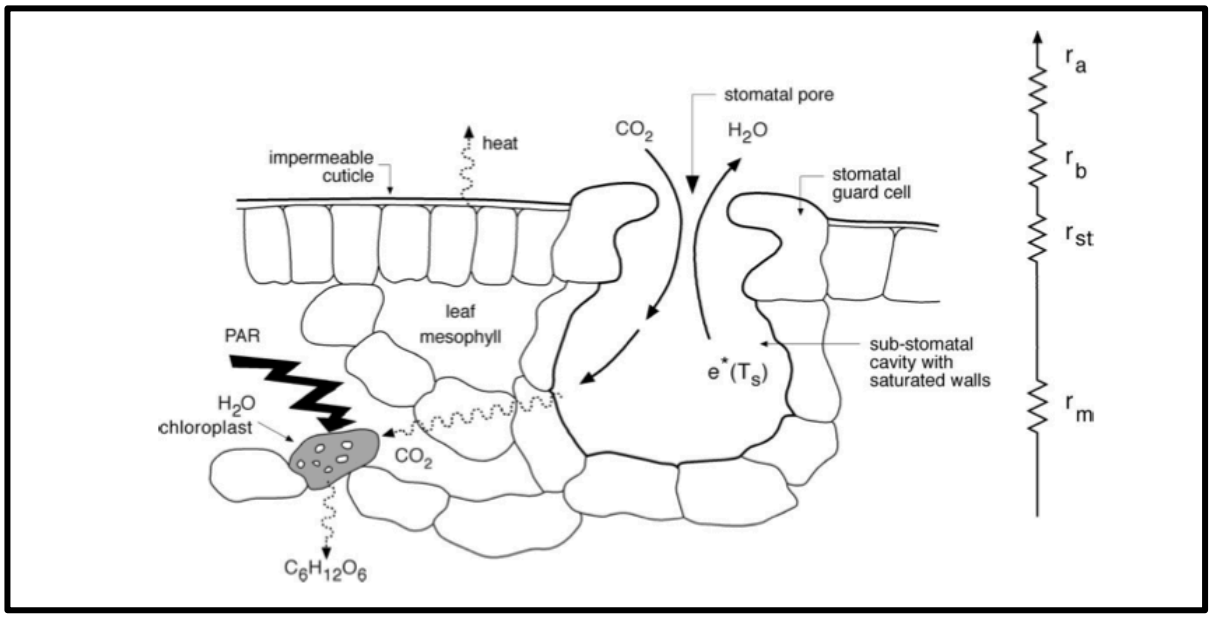


Figure 0.3: Cross section of a leaf (Pitman, 2003)

Canopy temperature can be used to measure the water stress of a crop and is easily obtained. The stomata of the plant closes when there is a shortage in terms of the amount of water available to the plant, this happens as a result of the plant having less turgor pressure. This leads to a decrease in the amount of transpiration and evaporation energy and will source a higher leaf temperature (Figure 2.3) (Van Zyl, 2017). Thermal cameras are used to measure the canopy temperature of plants on ground and aerial level (Gonzalez-Dugo *et al*., 2015).

2.7.4. Stomatal conduction

Stomatal conductance is the rate of the carbon dioxide diffused into the leaf, or the water vapour leaving the leaf through the stomata during photosynthesis (Figure 2.3). A leaf porometer calculates the stomatal conductance to water vapour ($mmol.m^{-2}.s$) (Nardini *et al*., 2014).

There is a correlation between canopy temperature and stomatal conductance since stomatal conductance has a direct influence on transpiration cooling. A lot of the same environmental and physiological factors has an effect on both of these traits. A single measurement of canopy temperature can calculate a precise transpiration rate and stomatal conductance of the area. The water content and energy status of the crop are the two key components that determine the water status of the crop. Water potential is created by combining pressure, osmotic and matric pressure at a specific time. The hydration status of the tissue is specified by the relative water content of the crop. The relative water content is a cost-effective measurement that displays the presentation of genetic variability and intermediate heritability in wheat (Cossani *et al*., 2012).

Root systems define the volume potential of soil which is the amount of water and nutrients the root system can use. The interaction between environmental and genetic factors has an influence on the effective root system, where one of the main traits necessary for cereal breeding is deeper and vigorous root systems that have advanced productivity of water in conditions with little water. Osmotic adjustment uses the build-up of the minor particles in the solution inside the cells which will result in a decrease in water and will aid in maintaining the function of the cell in water stress conditions. It has been classified as a mechanism to maintain grain yield under stressed conditions

through maintenance of nutrients and water as well as growth of roots (Lopes & Reynolds, 2012).

2.8. Different selection methods

2.8.1. Pedigree selection

Pedigree selection comprises of the selection of different genotypes based on their ancestral reproductive ability and the mixture of genes that are desired and are present. It is therefore crucial for the breeder to look at the pedigree of the cultivar and to keep a good record of their performance. This will particularly be beneficial if both of the parents have desirable phenotypes such as resistance to diseases, high yield and flexibility. Pedigree selection will start in the second generation, since the progeny will be heterozygous. The best plants within the optimum families that express the most wanted phenotypes are selected. The homozygosity will increase with each generation and after the fourth-generation selection will be ineffective. This selection method is suitable for the selection of lines from segregating plants, since it compares both the phenotypes and genotypes. This method is ineffective when the aim is to accumulate the number of minor genes necessary to provide horizontal resistance (Ali, 2011).

2.8.2. Bulk population breeding

Bulk population breeding is a method where the change in gene frequencies depends on natural selection that most likely will lead to crop improvement. This method most likely leads to the development of lines that are superior, since they are more likely to reproduce, survive and be more productive. It requires that the F_1-generation are cultivated in bulk in order to harvest sufficient seeds for the F_2-generation. This bulk cultivation will be repeated until the F_4-generation have been harvest or the homozygosity are sufficient. A negative aspect of this method is that the desired genotypes may lessen, and undesirable genotypes is increase (Tee & Qualset, 1975) (Wang *et al*., 2003).

2.8.3. Single seed decent

Single seed decent (SSD) was first suggested by Goulden (1941) but improved by Brim (1960). The SSD technique is almost similar to the pedigree method, but with a few adjustments that will increase the tempo. This technique entails the selection of parents, which are crossed, to develop a base population. This step is followed by the random selection of seeds from every plant and then to sample it in the following generations. This technique will aid to decrease the number of generations in order to get homozygosity without selection. The downfall of this method is that selection is based on individual phenotypes and not on the performance of progeny (Pignone *et al*., 2015)

2.9. Remote sensing in breeding programs

The phenotype of wheat is a combination of physiological, morphological, structural and performance related traits of a particular genotype in a determined environment. A phenotype is the result of the influence of biotic and abiotic factors on the wheat's genes. Remote sensing is the measurement of an item by a device without any physical contact between them that can be used to do high throughput phenotyping that can contribute to making decisions, yield predictions and forecast spatial field variability (Karytzis *et al*., 2017). Remote sensing can be used to make detailed observations of crops in a specific and non-invasive manner. Thorough examining of crops in the field will help to identify novel traits, quantify traits that will aid in the performance of crops and to analyse genetic control of complex traits (Virlet *et al.,* 2017).

The foundation of spectral science and remote sensing is the capability to quantify electromagnetic energy at different wavelengths associated with parts of the plant. Spectral imaging is used to quantitatively measure phenotypes through the interaction between plants and lights, such as absorbed, transmitted, reflected and emitted photons. A healthy and diseased plant responds differently with electromagnetic radiation and this will help to distinguish between them (Montesinos-López *et al.,* 2017).

2.9.1. The Red Green Blue colour model

The most straightforward and recognized colour model is the Red Green Blue (RGB) as well as its subsection Cyan Magenta Yellow (CMY) colour model and is the closest similarity to how humans perceive colour. This model also has a resemblance to the principles of additive and subtractive colours. Additive colours are generated by mixing spectral light in different groupings of solid spectral colours that are optically combined by being positioned narrowly together. This may cause that two or more colours being observed as one singular colour. Subtractive colours are observed when pigments in an item absorb specific wavelengths of white light while reflecting the rest (Kim *et al*., 2015). The human eye has evolved to perceive light at wavelengths between 380 and 780 nm, nevertheless this range is a part of the electromagnetic spectrum known as “visible light”. The three RGB wavelengths respectfully are red (700-635 nm), green (560-520 nm) and blue (490-450 nm) in the visible light spectrum (Hsu *et al*., 2015).

Leaf area index (LAI) and biomass are two variables that are linked to grain yield in wheat and are a good indicators of plant health. It is, therefore, necessary to measure these variables during wheat development in order to make predictions as well as selections. Through carrying it on a Remote Pilot Aircraft System (RPAS), commercial RGB cameras can be used to estimate these growth status variables. The canopy of the plants can be observed from the aerial view and the volume can be calculated (Fan *et al*., 2017).

2.9.2. Spectral vegetation indices

The estimation of spectral vegetation indices (SVI) is the most popular technique used to retrieve information about crops through remote sensing. This is mainly centred on formulas that use the light that reflect of the canopy at diverse wavelengths that are within the near infrared electromagnetic spectrum. There are a variety of SVI used such as normalized difference vegetation index (NDVI), the Simple Ratio (SR) and green normalized difference vegetation index (GNDVI) (Karytzis *et al*., 2017).

Simple ratio is a ratio-based index or ratio vegetation index (RVI). It is calculated by dividing the near-infreared (NIR) by the red wavelength. It displays a low value for soil,

ice and water and a high value for vegetation. This index also reduces the effects of topography and the atmosphere. GNDVI is associated with grain yield when documented at booting and anthesis and explains more variability that NDVI and SR and has a superior discriminating efficiency between the wavelengths and can be an enhanced predictor of grain yield when it is documented at an early stage (Karytzis *et al*., 2017).

NDVI has shown significant correlation with grain yield (Sun *et al*., 2017). In a study done by Karytzis *et al*. (2017) there was a significantly positive correlation between SVIs with SPAD values and photosynthetic pigments and this endorses the close associations between canopy greenness and SVIs. There was also a significant correlation between chlorophyll a and NDVI/SR. There was a significant positive correlation between SPAD values with SVIs at both milk and dough stages. The difference between the two wavelength bands can be used to generate different indices in the vegetation of crops.

NDVI values range between -1 to +1, where a negative value indicates a high possibility of water while a value close to +1 indicates green leaves. When the value of the NDVI is close to 0 there is no green leaves and is most likely an area with no plants. Plants with healthy vegetation and chlorophyll content absorbs more blue and red light and reflects more green and NIR light than a stressed plant (Figure 2.4). This is the reason why we visualize vegetation as the colour green and if NIR could be seen it would be seen for vegetation as well (Gandhi *et al*., 2015).

Remote-sensing of canopy temperature can be done by infrared thermometers and can be used to estimate the water-stress of the wheat. In the case of wheat experiencing water-stress wheat genotypes with lower canopy temperatures during midday has a better plant water-status. In a wheat breeding program, canopy temperature is, therefore, used as a selection index for resistance against water shortage (Blum *et al*., 1989). Canopy architecture and development (e.g. leaf size and angle, plant height, emerged spikes and soil covering) should be considered, when using canopy temperature for breeding line selection. Multi-spectral imagery allows for the evaluation of complex traits, for example canopy photosynthesis and fluorescence under natural sunlight conditions (Araus & Cairns, 2014). Canopy temperature and

Figure 0.4: Comparison of how light reflects on leaves under different conditions (Gandhi *et al*., 2015).

NDVI can be used to estimate wheat yield even before harvesting. Spectral vegetative indices (VIs) are used to evaluate the physiological condition and agronomic traits of crops. The indices used are dependent on the specific species due to different species having dissimilar leaf structures and canopy architectures (Montesinos-López *et al.,* 2017). Chlorophyll content meters have been commercially used for more than 25 years and provide rapid and accurate non-destructive chlorophyll analyses of leaves (Raymond *et al*., 2014).

Ground-based remote sensors like sensors on tractors and hand-held sensors have high resolution for phenotyping studies. They are, however, limited mostly due to the time required, the scale they can be used for and the limitations on their portability (Araus & Cairns, 2014). During the development and growth of a plant the soil plays a major role and it is challenging to simulate it. The features of the environments are not understood and monitored, which makes it difficult to mimic the field environments in controlled environments. Phenotyping in the field conditions causes a genetic bottleneck for future breeding advances (Araus & Cairns, 2014).

It has been reported that the use of satellite platforms is an appropriate method to evaluate LAI as a spatial and temporal assessment. Some studies overestimated LAI up to 17% as a result of the understorey element from satellite imagery. This imagery method has restricted application in breeding programs due to the platform's low temporal and spatial resolution. The resolution of these imagery platforms has been improved in the latest commercial satellites for example Worldview (1 and 2), RapidEye, Ikonos, GeoEye and Pleyades. The images from these satellites are, nevertheless, costly and requires a lot of knowledge of how the technology and science works (Fuentes *et al*., 2014). Currently, free satellite images have a resolution of 10 m^2, which is not sufficient for studies where the resolution should be higher (Luo *et al*., 2018).

2.10. High-throughput phenotyping platforms in breeding programs

The use of high-throughput phenotyping platforms could be an economical solution for linking phenotype with genotype through enlarging the precision and capacity while decreasing the time to evaluate large experimental sites (Haghighattalab *et al*., 2016). These HTPPs must improve to quantitatively evaluate a large number of plant

phenotypes to predict how the plants will perform when cultivated. Remote Pilot Aircraft Systems (RPASs) can contribute to overcome this challenge due to its low cost and tool to examine factors such as limited recourses, population growth, and climate change and food security. A RPAS is a powered aerial vehicle that does not carry a human operator that uses aerodynamic forces to supply lift. This vehicle can fly by instructions from a remote pilot or autonomously and can carry a payload like a camera. It can also increase the rate of genetic improvement in wheat through increasing the capacity to evaluate large numbers of plants in field trials (Busemeyer *et al*., 2013).

RPAS phenotyping platforms make it possible to characterize many plots, overcoming the limitations linked to ground-based phenotyping platforms. Some breeding populations can comprise of 5000 lines and these platforms can be used to simultaneously evaluate all lines (Araus & Cairns, 2014). HTPPs are used to assess the phenotypes of plants in a quicker and more precise manner than the methods used traditionally. They are practical in breeding programs and must be amenable to field conditions, plot sizes and experimental designs. A large number of lines can, therefore, be evaluated at a lower cost, within a shorter time-period and with less manual labour than current methods. The use of information technologies such as efficient computational tools and proximal or remote sensing are required to characterize the growth responses of genetically different plants in the field and correlate them with the individual genes.

Different to airplanes, a RPAS can operate in a stationary position and capture high quality images. Most of these RPAS use RGB/colour infrared (CIR) cameras along with sensors for thermal imaging (Araus & Cairns, 2014). RPAS-based RGB imaging can be used for the plant height and RGB-based vegetation indices (VI_{RGB}) analysis, where both of these has a centimetre resolution and can be obtained from a multi-rotor RPAS. Visible band vegetation indices and their combination with plant height have potential for biomass estimation in early growth stages (Bendig *et al*., 2015).

Many studies have been done with the use of hand-held sensors. This equipment has several limitations, such as being labour intensive as well as time consuming. Low altitude aerial platforms can be used to overcome the limitations at a low cost and at a

high spatial and temporal resolution. Recent studies have shown that correlations between agronomic traits and SVIs obtained from aerial imagery are similar and, in some cases, better than ground-based measurements (Karytzis *et al*., 2017).

The technology of RPAS equipment and software has been improving quickly in the past years and the use of this platform is becoming more popular in both environmental and agricultural applications and studies. The materials that are used in these vehicles varies and materials such as carbon fibre is used in order to make it more robust, but still maintaining a lighter overall weight. The automation software of the RPAS has developed to be more user friendly and the programs to do the processing of images are more available and, in some cases, it is open software. The vehicles can be programmed to operate at different speeds, heights and routes to fit the requirements needed to capture the needed images and slight weather changes are usually not a problem for it (Simelli & Tsagaris, 2015).

Leaf area index and biomass are two major variables that are linked to grain yield in wheat and are good indicators of plant health. It is, therefore, necessary to measure these variables during the development of the wheat in order to make predictions as well as selections (Fan *et al*., 2017). The canopy of the plants can be observed from the aerial view and the percentage ground coverage can be calculated which can be linked to LAI. A RPAS with a RGB camera can also be used to process a digital surface and elevation model to provide a 3-dimentional model as well as the estimation of plant height. Plant height can also be a good indicator of the biomass of the plant (Madec *et al*., 2017).

Several studies make use of digital image analysis to determine chlorophyll content and nitrogen status based on leaf colour, however there are a lack of detailed methodology and software used. Most of these field measurements are executed on an image of the entire area that includes the soil and other obstacles. A key problem is that it is overlooking the key benefit of image analysis which is the exclusion of pixels that does not belong to the plants when the analysis is done. Another obstacle to overcome is to distinguish between the plants, weeds or other plants (Baresel *et al*., 2017).

2.11. Traits that influence wheat yield

Wheat grain yield is one of the most important factors for all wheat breeders and farmers and is influenced by a combination of certain physiological processes, genes as well as response to different environments (Simmonds *et al*., 2016). Factors influencing grain yield is divided into certain categories of traits that include plant height (Wu *et al*., 2012), spike length (Zikhali & Griffiths, 2015), spike number (Sharma *et al*., 2003), grain number (Abdullah *et al*., 2015), tiller number (Kumar *et al*., 2015), grain weight, thousand kernel weight (Abdullah *et al*., 2015), grain size, grain density and physical parameters that include grain width, length and area. It is important to select for these traits to improve the yield of the wheat grain (Gegas *et al*., 2010).

2.11.1. Plant height

The height of wheat plants is a crucial trait, where a reduced plant height is more likely to have a higher yield than taller plants (Würschum *et al*., 2017). Taller wheat plants do not have sufficient stems to support the heavy grains and consequently leads to the lodging (falling over) of plants. The effect of this problem was reduced by the introduction of the semi-dwarf genes from the cultivar “Norin 10”, which resulted in a lower plant height and less lodging that consequently lead to higher yield (Hedden, 2003). This specific cultivar, which was developed by the Japanese early in the 1900s, contained the two dwarfing alleles, *Rht-B1b* and *Rht-D1b*. The American agronomist, Norman Borlaug, received one cross that he used to breed cultivars that could adapt and reproduce within tropical and sub-tropical climates. The progeny of the cultivar was dispersed to numerous countries and there was a tremendous increase in wheat yields (Hedden, 2003). The lower plant heights of the semi-dwarf cultivars also lead to an enhanced harvest index. The harvest index for the majority of cultivars were 0.3 as a result of the losses in grain yield due to the plants lodging, but the harvest index for the cultivars that contained the semi-dwarfing genes was 0.5 because of the shorter plants (Sakamoto & Matsuoka, 2004). These semi-dwarf genes are still being used around the world for the reduction of plant height (Würschum *et al*., 2018).

2.11.2. Spikelet number and spike length

Spike length is an important trait to select for when increasing the grain yield, since a larger spike will develop more grains than a spike with a smaller length. The spike can also be capable of chlorophyll retention that will allow them to stay functional and undergo photosynthesis for an extended period which includes the awns as well. These features will contribute to the increase of dry matter in the kernel by approximately 20-30%. Therefore, an alternative method to increase grain yield will be to select and breed for longer spikes that will lead to the development of an increase in florets and spikelets (Sharma *et al*., 2003).

The spikelet number per spike is also an important component for the increase in grain yield since it is linked to an increased grain number per spike. The earliness per se (*Eps*) gene, *Eps-Am1* that is located on chromosome 1AL of *T*. *gonococcus*, is reported to regulate the spikelet number and is also connected to the amount of grain per spike (Zikhali & Griffiths, 2015).

2.11.3. Grain number and weight

Grain yield in usually defined as the combination of the grain number and the overall grain weight, while these two factors can easily be measured and is also distinguishable from each other. An increase in the amount of grain yield has a positive correlation with a higher amount in grain number. Traits that can be associated with the enhancement of the yield was mostly measured during the growth stage of wheat, though currently yield is mainly determined on traits that is measured at the anthesis growth stage of the plant's life cycle (Abdullah *et al*., 2015).

Grain number per ear relies on the development of the florets, their ability to survive inside each spikelet and the state of fertilization during anthesis. Only a proportion of the total amount of florets ultimately develops into grain (Siddique *et al*., 1989). Since there are about 16-18 spikelets and there can be as much as 11 florets in each spikelet where a maximum of 40 grains eventually develop, the amount of potential sites surpasses the amount of grains that develop (Sinclair & Jamieson, 2006).

The relationship between grain weight and the amount of grain has negative correlation, because a higher grain number is mostly connected to the spikelets having more grains that most likely have a lower grain weight. When a plant has a higher grain weight, the grain number will be less which will lead to a lower yield potential. This indicates that breeders should rather aim to maintain grain weight instead of trying to increase it (Fischer, 2011).

Thousand kernel weight (TKW) is an important agronomical trait that is mainly used to represent grain weight and has a positive correlation with wheat yield along with grain number. TKW is, therefore, one of the traits breeders put the most emphasis on as it is one of the most stable mechanisms of grain yield. However, this trait is quantitatively measured, and environmental factors can influence this (Groos *et al*., 2003). Kumar *et al*. (2015) discovered 10 Quantitative Trait Loci (QTLs) for TKW that are located on 8 different chromosomes. The main QTLs was QTKW.ndsu.5A.1 and QKW.ndsu.5A.2 that can be found on chromosome 5A (Kumar *et al*., 2015).

2.11.4. Grain size

The length, width and area of the grains have a positive correlation with grain yield. The *TaGS5-3A* (Ma *et al*., 2015) and *TaGW2-6B* (Qin *et al*., 2014) genes are a positive regulator for grain width and weight and their presents can be linked to an increased yield. The *TaGS-D1* gene is correlated with having in increased length and weight (Zhang *et al*., 2014). Some machines and mobile applications are used to measure the grain size by image-based analysis. The SeedCounter mobile application is a good example of grain size measurements (Komyshev *et al*., 2018).

2.11.5. Tiller number

Tiller development is one of the important traits that have an influence on wheat structure and yield. Dwarfing genes of the semi dwarf wheat's developed more superior tillers per plant as well as increased biomass leading to a modification of the yield potential. Tillers that originate from the wheat's stem are known as the primary tillers and those that arise from the primary tillers are known as secondary tillers. The survival capacity and the appearance of the tillers are mainly responsible for the tiller number.

Fertile tillers, therefore, contributes to grain yield (Kumar *et al.,* 2015). The tillers develop from the outgrowth of axillary buds that are found within the axils of the leaf on the basal internodes (Longnecker *et al*., 1993).

Spielmeyer & Richards (2004) discovered a molecular marker for tiller number. They identified the tiller inhibition gene (*tin*) that is responsible for a decrease in tillering and is also associated with larger spikes, a decrease in sterile numbers and an increase in grain size as well as harvest index. As a result of the agronomic potential of the *tin* gene a microsatellite marker, Xgwm136, was developed that is associated with the gene and was mapped on the chromosome 1AS. It was, therefore, recommended that this marker be used for MAS in order to detect low-tillering lines.

2.11.6. Days to heading

The amount of days to heading trait is important and is measured in cereal crops because cultivars with the most appropriate heading time within the life cycle of the plant and the environment, will lead to a higher wheat yield potential within different environments (Del Pozo *et al*., 2016). A crop that has a sensitive photoperiod will not pass the vegetative phase until it has reached the required increase in day length or the photoperiod for that crop. Conversely, when there is an increase in temperatures, a crop that is not sensitive to the photoperiod will instantaneously change from the vegetative phase to the reproductive phase and the crops will develop and complete grain filling prior to the high temperatures in the summer season (Rahman *et al*., 2009, Iqbal *et al*., 2011).

Wheat cultivars are different when it comes to the particular developmental stages and anthesis and it is mainly controlled by a collection of genes that affect necessity for vernalisation (*Vrn* genes), sensitivity to photoperiod (*Ppd* genes) and earliness per se (*Eps* genes). The homologous genes (*Vrn-A1*, *Vrn-B1* and *Vrn-D1*), that are located on the 5A, 5B and 5D chromosomes, are mainly responsible for controlling the vernalisation of the crops. *Ppd-A1*, *Ppd-B1* and *Ppd-D1*, situated on group 2 chromosomes, are mainly responsible for regulating the sensitivity to photoperiod. The heritability of *EPS* is uncertain but includes the locus *Eps-2B*, which is situated on chromosome 2B and other loci on the chromosomes 3A, 4A, 4B and 6B. The result of

vernalisation as well as photoperiod will possibly be different with the developmental stage and additional genes may influence the response to temperatures during development (Herndl *et al*., 2008).

2.11.7. Wheat harvest index

The term harvest index (HI) is used to describe and quantify in ratio the proportion of harvested grain to that of the total aboveground crop biomass (HI=Grain yield/Total above ground biomass at maturity). It is a measurement of the relative efficiency with which a plant moves assimilated materials from vegetative tissue such as leaves, stems and branches into the developing and mature grain. It is a suitable way to describe the efficiency with which plants make use of sunlight and raw materials (water and nutrients) from the soil to increase the grain yield (Dai *et al*., 2016). Foulkes *et al*. (2009) proposed 0.60 to be the theoretical limit for a harvest index, while the majority of the values varied between 0.4 and 0.5. Furthermore, this leaves room for improvement as a harvest index of 0.56 was accomplished (Foulkes *et al*. 2009).

2.11.8. Flower fertility

Each of wheat spikelet has the potential to develop eight florets. The final number of grains per spikelet at the physiological maturity growth stage has a major influence on the grain yield. The floral structures, however, has to endure an abortion and development process before the final grain number can be measured. The yield potential can be measured only after the maximum floret primordia amount is reached. The fertility of the flowers at the anthesis stage is determined by the floral degradation process, while one to three florets are lost during post-anthesis events until the final grain amount can be measured at the maturity growth stage (Guo & Schnurbusch, 2015).

Chapter 3: Materials and methods

The need to increase wheat yield is an important factor for producers and breeders, but there are several challenges. The aim of this study was to identify germplasm with high-yielding traits and to introduce them into the MS-MARS facilitated pre-breeding program of the SU-PBL by applying crossbreeding. This was achieved by following four objectives.

The first objective was to identify germplasm that can benefit the breeding program. This germplasm should have genes that are linked to high-yielding physiological traits as well as other factors such as baking quality and rust resistance genes. DNA were extracted and molecular markers along with microsatellite markers were used for screening of these genes. The results were analysed, and donor material was selected.

The second objective was to introduce the germplasm to the MS-MARS facilitated pre-breeding program of the SU-PBLs population with the aim to increase the yield. A selection of high-yielding lines was used as pollen donors to cross with the male sterile population. The progeny seeds of these cross-bred seeds were used in the next cycle and two cycles was accomplished.

The third objective was to apply HTTPs to make detailed and non-invasive observations of wheat in an individual manner without being invasive. Remote sensing was done by the use of a RPAS, with different cameras and sensors over two separate seasons. A commercial and a designed in-house pipeline was used to create orthomosiac of the images and perform the photogrammetry analysis separately. Traditional hand-held instruments to measure the LAI, stomatal conductance, CCI and temperature of each line was also done over both seasons and was compared to the HTTPs in order to develop a prediction model.

The fourth objective was to perform post-harvest phenotyping on the high-yielding lines, where the following traits were measured: days to heading, grain weight, thousand kernel weight, harvest index, plant height, tiller number, grain/spike, flower fertility, spikelet number, spike length, grain area and length. These several traits have been identified to be related to an increased wheat yield. Additional traits were

measured on the NIR grain analyser. Statistical analysis was performed to identify lines with high yield-related traits.

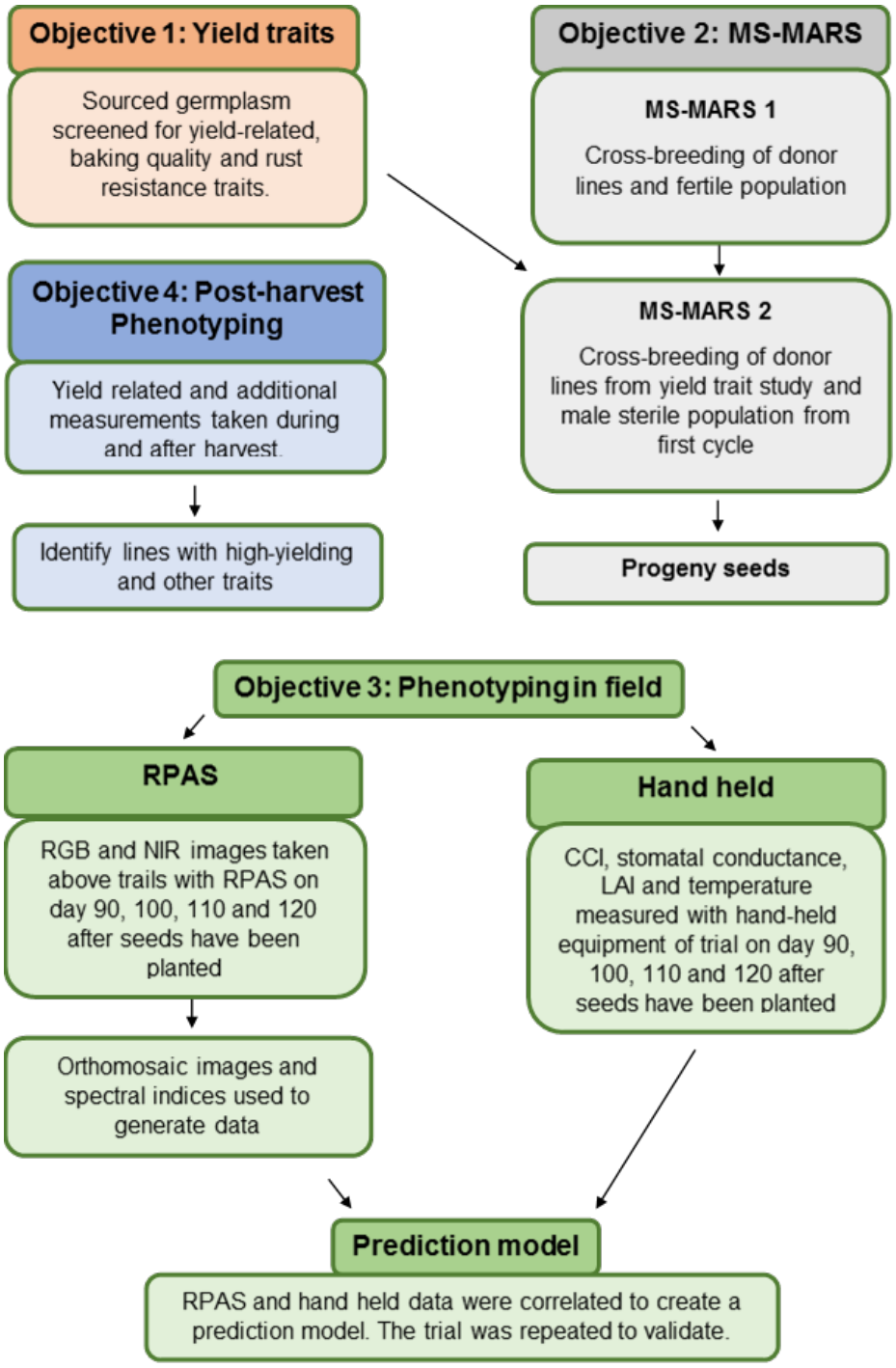


Figure 0.1: Flowchart of objectives for this study

*Male sterility marker assisted recurrent selection scheme (MS-MARS), red green blue (RGB), near infrared (NIR), chlorophyll content index (CCI), leaf area index (LAI), remote pilot aircraft system (RPAS).

3.1. Plant material for pre-breeding

Thirty possible donor lines from the 5th Wheat Yield Consortium Yield Trial (WYCYT) nursery was received, with a list of the combined analysis results under two different sowing dates in Obregon Mexico during 2017. Three additional genotypes from CIMMYT, 'Bluebird' (Christopher *et al*., 2013), 'Roelfs' and 'Atilla' (Aisawi *et al*., 2015) with high chlorophyll retention and photosynthetic capacity were also included.

3.2. Donor material selection

Prior to molecular characterisation, 22 wheat lines were selected based on their performance in 2017. The performance traits selected for were (1) TKW, (2) HI, (3) grain number per m^2, (4) spikes per m^2, (5) grain number per spike and (6) if the plant height was optimal. Genotypes that had a trait value larger than the mean for all genotypes for a minimum of one trait were selected. The 22 lines that were selected for molecular characterisation were planted in foam trays containing soil and grown in a glass house. Duplicates were planted for each of the wheat lines. The plants were watered by hand daily and cut when they reached a three-leaf stage. Leaves from the plants in each foam tray were pooled and stored in a 2 ml micro centrifuge tube.

3.2.1. Protocol for CTAB extraction

DNA extraction was performed by employing an adapted shortened version of the method described by Doyle and Doyle (1990). Approximately 100 mg of plant leaf tissue was cut into smaller pieces and placed into a 2 ml micro centrifuge tube together with three 3 mm stainless steel beads. 500 µL 2% (w/v) cetyltrimethylammoniumbromide (CTAB) [20 mM EDTA (pH 8) and 1.4 M NaCl, 100 mM Tris-Cl (pH 8)] extraction buffer was also added to the 2 ml microcentrifuge tube. The plant material was grinded with a Qiagen® Tissue Lyser (Qiagen, Southern Cross Biotech, Claremont, RSA) at 32 Hz for three rounds of 30 s. The plant mixture was incubated in a water bath for 30 min at 60°C. After incubation, 500 µL of chloroform:isoamylalcohol (24:1) was added to each tube and inverted three times, followed by centrifugation at 14 000 rpm for eight min. The supernatant was transferred into a sterile 1.5 ml microcentrifuge tube, 400 µL of chloroform:isoamylalcohol (24:1) was added and the tube was mixed by inversion before it was centrifuged for 5 min at 14 000 rpm. The supernatant was transferred into a sterile 1.5 ml microcentrifuge tube

and 50 µL of 3M NaOAc was added along with 500 µL of 100% ethanol that was stored at -20°C. The microcentrifuge tube was then carefully inverted several times for DNA precipitation to occur. The precipitant was collected through centrifugation at 14 000 rpm for 2 min, and the subsequent supernatant was discarded. The DNA was washed by adding 500 µL of 70% ethanol and centrifugation for 2 min at 14 000 rpm. The supernatant was discarded, and the step was repeated. After the washing step, the pellet was air-dried and re-suspended with 30 µL of distilled water (dH_2O) and stored at -20°C. The quality and quantity of the extracted DNA was measured using a Nanodrop® ND-1000 spectrophotometer (Thermo Fisher Scientific Inc., Kempton Park, RSA). The DNA was diluted with DNase/RNase-free water to a concentration of 100 ng/µL. The stock was stored at -20ºC while the diluted solution was stored at 4ºC.

3.2.2. Genotyping of the selected plant material

The SU-PBL employs a standardised panel of molecular markers (Table 3.1) that are routinely used for screening of different wheat nurseries. This panel includes markers for rust resistant genes, reduced height genes for optimal height of the plants and gluten genes, which has an influence on the baking quality of the wheat. The SU-PBL additionally use a group of six microsatellite markers that are used as a minimum marker set for the evaluation of the genetic diversity of wheat lines and material to select for backcrossing in the nurseries (Honing, 2007).

The PCR reactions, for marker analysis, of all the samples were performed using a MyCyclerTM Thermal Cycler (Bio-rad, USA), Thermal Cycler (Applied Biosystems, Fairlands, RSA) or a TECHNE TC-5000 (Lasec, Cape Town, RSA). The primers for each of the reactions had a stock concentration of 10 µM and was developed by Integrated DNA Technologies (Whitehead Scientific Inc, Stikland, RSA) in addition to the KAPA Green Readymix that was bought at KapaBiosystems (distributed by Lasec SA (Pty) Ltd, Cape Town, RSA) along with the One Taq Quick-Load 2x Master Mix that comes with the necessary Standard buffer from New England BioLabs (distributed by Inqaba Biotechnical Industries (Pty) Ltd).

All molecular markers were dominant, therefore when the amplicon of the desired allele was visible it was documented as 1; subsequently when the undesirable allele was

visible it was documented as 0 to indicate the absence of the desired allele. The allele frequency was then calculated by dividing the sum of the plants that had the desired allele by the total of all the plants.

Table 0.1: SU-PBL standard panel of molecular markers

Trait	Genes	Primers	Primer sequences	Ta (°C)	Fragment size (bp)	Citation
Disease resistance genes	*Sr2* (positive)	Xgmw533-F	5'-AAGGCGAATCAACAGGAATA-3'	62	120	Spielmeyer *et al.* (2003)
		Xgmw533-R	5'-GTTGCTTTAGGGGAAAAGCC-3'			
	Sr2(negative)	X3B028-F	5'-ACGAACAAGGGGAAGACG-3'	62	243	McNeil *et al.* (2008)
		X3B028-R	5'-TTTCGGTAGTTGGGGATGC-3'			
	Sr2 (CAPS)	CSSr2-F	5'-CAAGGGTTGCTAGGATTGGAAAAC-3'	60	53, 112, 172	Mago *et al*., (2011)
		CSSr2-R	5'-AGATAACTCTTATGATCTTACATTTTTCTG-3'			
	*Sr24/Lr24**	SCS719-F	5'-TCGTCCAGATCAGAATGTG-3'	55	719	Cherukuri *et al*. (2003)
		SCS719-R	5'-CTCGTCGATTAGCAGTGAG-3'			
	*Sr26**	Sr26#43-F	5'-AATCGTCCACATTGGCTTCT-3'	60	207	Mago *et al*. (2005)
		Sr26#43-R	5'-CGCAACAAAATCATGCACTA-3'			
	*Sr31**	Iqg95-F	5'-CTCTGTGGATAGTTACTTGATCGA-3'	55	1030	Mago *et al*. (2005)
		Iqg95-R	5'-CCTAGAACATGCATGGCTGTTACA-3'			
	*Lr19**	12C-F	5'-CATCCTTGGGGACCTC-3'	60	119	Prins *et al.* (2001)
		12C-R	5'-CCAGCTCGCATACATCCA-3'			
	*Lr37/Sr38/Yr17**	VENTRUIP	5'-AGGGGCTACTGACCAAGGCT-3'	65	259	Helguera *et al*. (2003)
		LN2	5'-TGCAGCTACAGCAGTATGTACACAAAA-3'			
	*Lr34/Yr18/Pm38**	L34DINT9-F	5'-TTGATGAAACCAGTTTTTTTTCTA-3'	58	517	Lagudah *et al*. (2006)
		L34PLUS-R	5'-GCCATTTAACATAATCATGATGGA-3'			
Dwarfing genes	*Rht-B1b*	BF	5′-GGTAGGGAGGCGAGAGGCGAG-3′	58	237	Ellis *et al*. (2002)
		MR1	5′-CATCCCCATGGCCATCTCGAGCTA-3′			
	Rht-D1b	DF	5′-CGCGCAATTATTGGCCAGAGATAG-3′	58	254	
		MR2	5′-CCCCATGGCCATCTCGAGCTGCTA-3′			
Gluten genes	*Glu-Dx5*	P1	5'-GCCTAGCAACCTTCACAATC-3'	63	450	Ahmad (2000)
		P2	5'-GAAACCTGCTGCGGACAAG-3'			
	Glu-Dy10/	P3	5'-GTTGGCCGGTCGGCTGCCATG-3'	63	576/612	Ahmad (2000)
	Glu-Dy12	P4	5'-TGGAGAAGTTGGATAGTACC-3'			
Yield-related traits	*TaGW2*	TaGW2-6B-CAPS -F	5'-GGAAACAGAAATATCGCGG-3'	62	863	Qin et al ., 2014
		TaGW2-6B-CAPS -F	5'-AGGACTGTGGGGAATGAATG-3'			
	TaGS5	TaGS5-3A-CAPS -F	5'-AGCTCTGCTTCACGAGGAAG-3'	52	562	Ma et al ., 2016
		TaGS5-3A-CAPS -R	5'-CTCCTCTTTATATCGCGTCCC-3'			
	Ppd-D1	Ppd-D1 -1F	5'-TTCCCATAACTAAAACCGCG-3'	54	288	Wilhelm et al ., 2013
		Ppd-D1 -1R	5'-GGAACATCATTTCTGGACTTTG-3'			
	TaGS-D1	GS7D -F	5'-AGCGCTCGAAAAGTCAG-3'	56	522	Zhang et al ., 2014b
		GS7D -R	5'-GGCAGGTCCAACTCCAG-3'			

*Genes analysed in one multiplex reaction

**Annealing temperature (Ta), base pairs (bp),

3.2.3. Screening for rust resistance genes

The molecular markers for the rust genes *Sr2, Sr24, Lr37, Lr19, Lr34, Sr31* and *Sr26* were used to screen the selected male donors that will be introduced in the MS-MARS scheme if they displayed desired results. These seven molecular markers are part of a standard panel of markers that is routinely used within the SU-PBL (Wessels & Botes, 2014).

The PCR reactions for the *Sr2* molecular marker contained 7.5 µL of KAPA Green Readymix, 0.35 µM of CSSr2 forward and reverse primer, 1290 ng DNA and was brought to a final volume of 12.9 µL with dH_2O. The cycle conditions for the reaction consisted of a denaturation step at 94°C for 5 min, followed by 40 cycles of 92°C for 30 s, 60°C for 40 s and 72°C for 50 s and a final extension of 72°C for 5 min. An initial visualization step was done in order to confirm which samples needed to be restriction enzyme digested. Each of the PCR amplicons was visualised on a 1.0% (w/v) agarose gel. All the samples that had a visible amplicon of 337 bp were selected and were digested by the enzyme. The final reaction volume for the restriction enzyme digest was 2.5 µL and contained 1X Buffer O and 1 U/µL of PagI enzyme (Thermo Scientific) along with the remainder of the PCR amplicon. After incubation at 37°C for 20 min. The digested samples were visualised on a 2% (w/v) agarose gel stained with 4% (v/v) EtBr within a 1X TBE buffer at 100V. The positive control used for this reaction was DNA extracted from “Steenbras”, where the negative control used was extracted from “Chinese Spring”.

A multiplex-PCR was utilized to screen for the *Sr24*, *Lr37*, *Lr19* and *Lr34* rust genes. The final reactions volume was 17.8 µL. The reaction contained 12 µL KAPA Green Readymix along with 0.3 µM 719 forward and reverse primer, 0.3 µM LN2 primer and 0.3 µM VENTRUIP primer, 0.3 µM 12C forward and reverse primers, 0.36 µM L34DINT9 forward primer and L34PLUS reverse primer and 1.5 µL sample DNA. The cycling conditions for the reaction consisted of a denaturation step at 94°C for 5 min, a duration of 35 cycles with 94°C for 60 s, 57°C for 60 s and 72°C for 60 s and a final extension of 72°C for 7 min. PCR amplicons were visualised on a 1.8% (w/v) agarose gel. The positive controls used in this reaction was DNA extracted from “RL6064”, “W8417Lr19” and “Chinese spring”, along with a no-template negative control.

Another multiplex-PCR was utilized for the *Sr31* and *Sr26* rust genes that had a final volume of 18 µL. The sample contained 8.5 µL KAPA Green Readymix, 2 µL $MgCl_2$, 0.6 µM Iqg95 forward and reverse primers and 0.36 µM Sr26 forward and reverse primers along with 1.2 µL DNA. The cycling conditions consisted of a denaturation step at 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 60 s and a final extension of 72°C for 5 min. PCR amplicons were visualised on a 2.5% (w/v) agarose gel. DNA extracted from "Pavon31" and "Eagle" served as positive controls for the *Sr31* and *Sr26* genes respectively. A no-template control served as negative control.

3.2.4. Screening for dwarfing genes

Molecular screening for the two reduced height genes, *Rht D1b* and *Rht B1b*, were done separately. For the *Rht D1b* the final volume was 14 µL that consisted of 7.5 µL KAPA Green Readymix, and 0.54 µM DF forward and MR2 reverse primer as well as 1400 ng DNA and brought to volume with nuclease free water. The positive control was DNA extracted from the 'Pavon' cultivar and the negative was extracted from the 'Inia66' cultivar. The protocol for the reaction was a denaturation step at 94°C for 5 min, a duration of 30 cycles with 92°C for 30 s, 65°C for 30 s and 72°C for 80 s and a final extension of 72°C for 5min. For the *Rht B1b* gene 12.5 µL KAPA Green Readymix, 0.69 µM BF forward and MR1 reverse primers as well as 1570 ng DNA was used. Nuclease free water was used to bring it to volume of 15.7µL. The positive controls used was 'Inia 66' and the negative was 'Pavon'. The protocol for the reaction was a denaturation step at 94°C for 5 min, a duration of 30 cycles with 92°C for 30 s, 65°C for 30 s and 72°C for 80 s and a final extension of 72°C for 5min. Both the PCR products of the separate reduced height genes was electrophoresed on a 1.5% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V.

3.2.5. Screening for Gluten genes

Molecular screening of the gluten genes for determining the baking quality of wheat consisted of 12.5 µL KAPA2GTM Fast Multiplex PCR Mix and 0.38 µM P1 forward and P2 reverse primer for *Glu-Dx5*, 0.25 µM P3 forward and P4 reverse primer for *Glu-*

Dy10/Glu-Dy12 as well as 1250 ng DNA and was brought to volume of 20 µL. The positive control for *Glu-Dx5* was DNA extracted from "Chinese Spring" and the positive control for *Glu-Dy10/Glu-Dy12* was extracted from "Pavon31" with water as the negative control. The protocol for the reaction was a denaturation step at 94°C for 5 min, a duration of 30 cycles with 94°C for 30 s, 63°C for 30 s and 72°C for 30 s and a final extension of 72°C for 5min. 10 µL of the PCR products was electrophoresed on a 2.5% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V.

3.2.6. Screening for molecular markers associated with yield-determining traits

The PCR reaction for the yield-determining trait *Ppd-D1* molecular marker had a final volume of 15 µL that comprised of 8.5 µL of KAPA Green Readymix, 1500 ng DNA and the final concentrations of 0.33 µM of *Ppd-D1* forward and both reverse primers. The mixture was brought to final volume by adding nuclease-free water. The protocol for the reaction was a denaturation step at 94°C for 3 min, a duration of 35 cycles with 94°C for 30 s, 54°C for 60 s and 72°C for 60 s and a final extension of 72°C for 10 min. PCR products was electrophoresed on a 2.5% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V. The positive control used for this reaction was SNA extracted from the "Opata85" cultivar and the negative control was extracted from the "Chinese Spring" cultivar.

Each of the PCR reactions for the *TaGS-D1*, *TaGW2-6B* and *TaGS5-3A* molecular markers had a final volume of 10 µL that was comprised of 5 µL KAPA Green Readymix, 1000 ng of DNA, 0.25 uM of forward and reverse primers and was brought to volume with nuclease-free water. The protocol for the reaction for the *TaGS-D1* molecular marker was a denaturation step at 94°C for 3 s, a duration of 30 cycles with 94°C for 30 s, 52°C for 50 s and 72°C for 60 s and a final extension of 68°C for 5 min. PCR products was electrophoresed on a 3% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V. The positive controls used for this molecular marker was DNA extracted from the "Chinese spring", "Pavon F76" and "Opata85" cultivars while the negative controls were nuclease-free water.

The protocol for the PCR reactions of the *TaGW2-6B* and *TaGS5-3A* molecular markers included the touchdown method. The reaction was a denaturation step at

94°C for 30 s, a duration of 13 cycles with 94°C for 30 s, 62°C for 50 s and 68°C for 60 s, followed by 17 cycles with 94°C for 30 s, 55°C for 50 s and 68°C for 60 s and a final extension of 68°C for 5min. PCR products of *TaGW2-6B* were digested by the addition of 5 µL of the 0.5 U/µL BstNI enzyme (Thermo Scientific) that had a final volume of 10 µL and contained 1X NE Buffer and brought to volume with nuclease-free water. The product with the enzyme was incubated at 60°C for 20 mins. PCR products was electrophoresed on a 2.5% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V. The positive controls used for this molecular marker was DNA extracted from “Pavon F76” cultivar; and the negative controls used were extracted from “Inia F 66”, “Chinese Spring” and “Opata85” cultivars.

The enzyme digestion for *TaGS5-3A* had a final volume of 5 µL and contained 1X NE Buffer, 0.2 U/µL of Fnu4HI enzyme (Thermo Scientific) and brought to volume with nuclease-free water. The enzyme was added to the PCR product and incubated at 37°C for 30 min. The products were electrophoresed on a 2.5% (w/v) agarose gel stained with 4% (v/v) of EtBr within a 1X TBE buffer at 100V. The positive controls used for this molecular marker was DNA extracted from the “Inia F 66”, “Kite” and “Opata85” cultivars; and the negative controls used were extracted from “Pavon F76” and “Chinese Spring” cultivars.

3.2.7. Molecular diversity assessment

The genetic diversity of the selected plants was analysed by utilizing the set of six microsatellite markers that are routinely used in the SU-PBL (Honing, 2007). Each of the six primer sets, (Table 3.2) was used in single PCRs to screen for distinct microsatellite markers. Each reaction contained 6.25 µL KAPA Green Readymix, 10 uM forward and reverse primers, 1420 ng DNA and was brought to a final volume of 14.2 µL. The protocol for the reaction was a denaturation step at 94°C for 60 s, a duration of 45 cycles with 94°C for 60 s, 30 s at the annealing temperature and 72°C for 2 min and a final extension of 72°C for 10 min.

Table 0.2: Primer information on the SU-PBL set of microsatellite markers used for molecular diversity assessment

Primer name*	Primer Sequence	Ta (°C)**	Repeat sequence
Xgwm190-5D F	5'-GTG CTT GCT GAG CTA TGA GTC-3'	55	CT
Xgwm190-5D R	5'-GTG CCA CGT GGT ACC TTT G-3'		
Xgwm437-7D F	5'-GAT CAA GAC TTT TGT ATC TCT C-3'	47	CT
Xgwm437-7D R	5'-GAT GTC CAA CAG TTA GCT TA-3'		
Xgwm539-2D F	5'-CTG CTC TAA GAT TCA TGC AAC C-3'	60	GA
Xgwm539-2D R	5'-GAG GCT TGT GCC CTC TGT AG-3'		
Xwmc11-1A, 3A F	5'-CAC CCA GCC GTT ATA TAT GTT GA-3'	56	CT
Xwmc11-1A, 3A R 5	5'-GTT GTG ATC CTG GTT GTG TTG TGA-3'		
Xwmc59-1A, 6A F	5'-TCA TTC GTT GCA GAT ACA CCA C-3	58	CA
Xwmc59-1A, 6A R	5'-TCA ATG CCC TTG TTT CTG ACC T-3'		
Xwmc177-2A F	5'-AGG GCT CTC TTT AAT TCT TGC T-3'	52	CA
Xwmc177-2A R	5'-GGT CTA TCG TAA TCC ACC TGT A-3'		

*Table references: Xgwm (Röder *et al.,* 1998) and Xwmc (GrainGenes: A Database for Triticeae and Avena – Genetic Markers, [S.a.])

** Annealing temperature (Ta)

Polyacrylamide gel-electrophoresis (PAGE) was done in four main steps namely plate preparation, gel preparation, loading of the PCR samples and silver staining. During the plate preparation step 125 µL of plate glue was mixed with 25 ml of 100% ethanol. This was followed by adding 500 µL of the mixture to 1.5 ml 100% ethanol. 1.74 ml of the diluted plate glue and 140 µL 10% acetic acid were combined in a separate 2 ml microcentrifuge tube. Short and long glass plates were cleaned with 70% ethanol followed by wiping the short glass plate with the plate glue and the long glass plate with Wynn's C-thru. The plates were put together with a 1 mm spacer between them and were levelled with each other and clamped to ensure that they stay in that position.

For the preparation of the gel 40% acrylamide stock solution was made that consists of 0.129 M bis-acrylamide, 5.3 M acrylamide and was brought to 200 ml by adding distilled water. The stock solution was stored in the dark. A 6% sequencing gel solution was made up by combining 37.5 ml of the 40% acrylamide stock solution, 50 ml 5X TBE and 6 M urea. Eight hundred microliters of 10% ammonium persulfate (APS) solution and 160 µL N, N, N‟, N‟-Tetramethylethylenediamine (TEMED) was added to 160 ml of the 6% gel solution with. The solution was stirred and immediately casted on the plates.

The PCR samples were mixed with 10 µL loading dye (98% formamide, 10 mM EDTA, pH 8.0, 0.05% (w/v) bromo phenol blue, and 0.05% (w/v) xylene cyanol FF). The samples were denatured by incubating the tubes in a water bath for 5 min at 94°C before loading them onto the 6% polyacrylamide gels to undergo electrophoreses as described by Le Maitre (2010).

Silver staining was performed to visualize the PCR amplicons. The gels were placed in 2100 ml fixing solution (10.5 ml acedic acid, 210 ml 100% ethanol) for 20 min shanking, followed by rinsing the plates for 5 min with dH_20. It was then placed in 2100 ml staining solution (2.1 g AgNO3) for 20 min, followed by a 10 s rinse in dH_20 and placed in developing solution (8.505 ml formaldehyde, 31.5 ml NaOH) on the shaker until the fragments were clear and visible. The gel was subsequently rinsed in dH_20 and covered with plastic film to preserve the gel plates.

Data analyses for the microsatellite markers was done using PowerMarker v3.25 (Liu & Muse, 2005). The data were analysed as haplotype data and all of the default settings was used. The sub-dataset was used to calculate a statistical summary, which displayed the polymorphic information content (PIC) values and allele frequencies. The C.S. Chord distance matrix (Cavalli-Sforza & Edwards, 1967) was selected to compute the genetic distance between the wheat lines. Both neighbour joining (NJ) and unweighted pair group method with arithmetic mean (UPGMA) phylogenetic trees were created from the data and was viewed in MEGA version 5.

3.3. Male sterility marker-assisted recurrent selection scheme

3.3.1. Validation of the male sterility marker-assisted recurrent selection scheme

During this study, two cycles of the MS-MARS scheme were conducted over two consecutive years (Table 3.3).

3.3.2. Male sterility marker-assisted recurrent selection cycle one

Male fertile and sterile germplasm was obtained from the SUPBL's base population nursery established in 1999 (Botes, 2001), using the seeds harvested in the previous year's (2016) scheme. Both the male fertile and male sterile germplasm were planted

in separate greenhouses located on Welgevallen Experimental Station (WES) over a period of five weeks with three days between plantings. The four genotypes that had the best yield in the previous year's field trial were planted and represented the male population. Six seeds were planted per bag (125 × 105 × 230 mm) with daily water supply drippers for both the male sterile and male fertile populations.

Different male sterile and male fertile ears were phenotypically selected and were cut at the flowering phase in order to do cross-pollination. Since the base population was segregating 1:1 male sterile: male fertile plants, the male sterile plants were elected cautiously to avoid any selfing (self-pollination) of male fertile plants to facilitate hybridization. The tillers were arranged in the galvanized trays (60 cm x 45 cm x 16 cm) containing a nutrient solution to ensure optimal pollination. The trays were coated with antifungal paint. Trays were filled with 0,9% "Jik", 40 % nutrient solution and water. The nutrient solution consisted of 150 g of Microplex GA, 12.5 kg of SoluFert T3T and 100 L of water and 2: 6.8 kg of Calcinit mixed in 100 L of water. Fourteen litres of nutrient solution were added to a 1500 L tank with water. This central tray was developed with a nozzle in the front that acts as an inlet and a tap at the back that will be the outlet. The lids of the trays had the ability to facilitate 230 male sterile ears. The male fertile ears were placed in two separate trays that were also galvanized and had the ability to facilitate 70 male fertile ears. They were positioned on an elevated frame at 60 cm above the central tray. They were positioned at an angle into a galvanised trough that contained nutrient solution-soaked oasis® (floral foam) (Figure 3.2).

The male fertile ears were discarded after the male sterile ears were pollinated by them while the male sterile ears were left in the controlled nutrient solution for the period necessary for seed development and ripening. A pump to ensure air flow was used in the main trays to avoid fungal development. Plants were maintained in a growth room with light emitting diode (LED) lights that were set to optimize seed growth (Figure 3.2).

The nutrient solution was replaced every 14 days and the tips of the tillers were trimmed to ensure that the xylem and phloem were open, to avoid fungal growth and that there were no dead cells. Six weeks after fertilization, the tillers were removed from the troughs, placed in paper bags and dried in an oven at 21°C for seven days. Dried seeds were thrashed by hand and counted to assess the effectiveness of cross-

pollination. The seeds were then placed in paper envelopes and stored at room temperature (Figure 3.2).

Table 0.3: Male sterility marker-assisted recurrent selection cycles planting material and dates

Cycle	**Male fertile/sterile**	**Planting dates**	**Planting material**
MS-MARS cycle 1 (2017)	Male sterile	23, 30 March and 6,13 April	SU-PBL 2016 Nursery
	Male fertile	27 March, 3,10,17 April	HYLD field trial 2016
MS-MARS cycle 2 (2018)	Male sterile	16 May, 20 and 27 June	SU-PBL 2017 Nursery
	Male fertile	3, 10, 17 ,24 May	5th WYCYT nursery

* Male sterility marker-assisted recurrent selection (MS-MARS), Stellenbosch University Plant Breeding Lab (SU-PBL), High yield (HYLD), Wheat Yield Collaboration Yield Trial (WYCYT).

3.3.3. Male sterility marker-assisted recurrent selection cycle two

The second MS-MARS cycle was conducted in a similar way than the first cycle. Only four seeds, however, were planted in each of the pots for both the male fertile and male sterile populations. Seeds harvested from the first MS-MARS cycle served as the male sterile population while the male fertile population was the germplasm that were selected based on the results of the selection and genotyping step.

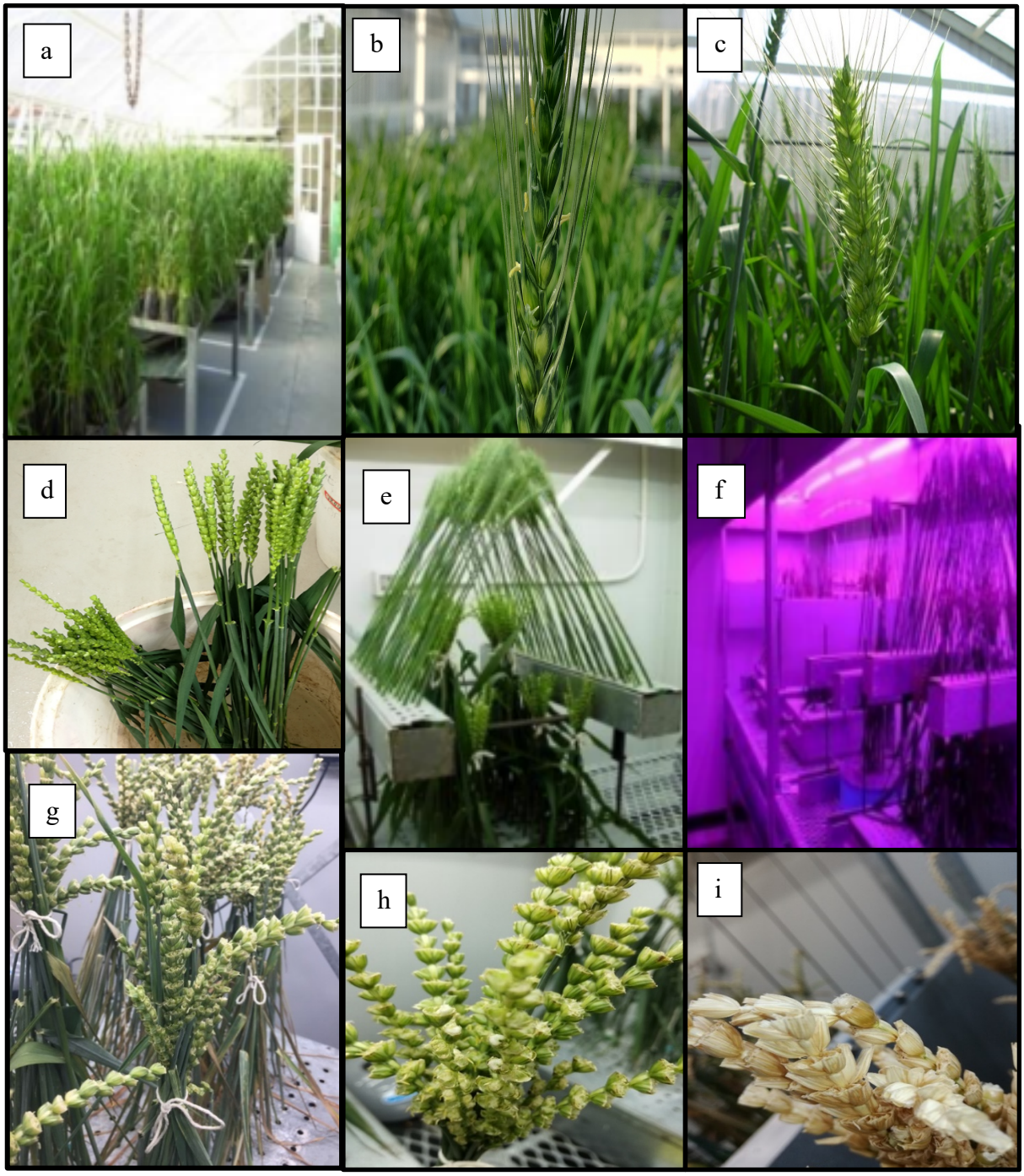


Figure 0.2: Male sterility marker-assisted recurrent selection cycle steps. a) Wheat plants are grown in the greenhouse. b) Male fertile tiller shedding pollen. c) The male sterile tiller ready for cross-pollination. d) The male sterile tillers cut open for cross pollination. e) The male fertile tillers arranged above the male sterile tillers. f) The light-emitting diode lightning in the growth room. g) Male sterile tillers receiving pollen from the male fertile tillers. h) Grain filling of the pollinated flowers. i) Seeds are ready to be placed in the oven and harvest.

3.4. 2017 season's phenotyping of high yielding genotypes

3.4.1. Field trial

The field trial wheat population that was used in this study was the high-yield (HYLD) trial of the SU-PBL and consisted out of 60 high-yielding wheat lines. Three replicates for each of the 60 entries were planted with a Wintersteiger Rowseed XL planter at Welgevallen Experimental Station (WES), Stellenbosch (-33.943752, 18.864628). A randomized complete block design (RCBD) created using AgroBase Generation II Version 18.3.1 (Agronomix Software Inc, Winnipeg, Canada), was used for the field layout.

3.4.2. Seed treatment

The 100 g seeds were placed in three different paper bags per entry. To treat seeds for the control of Stinking Smut (Bunt) (*Tilletia foetida*) and Loose Smut (*Ustilago tritici*), 1 ml of Anchor® Red (Arysta LifeScience South Africa) was added to each bag and shaken until dry.

3.4.3. Field Preparation and planting

Field preparation started after sufficient rainfall. The field was ploughed using a Claas 230 Talos tractor, followed by fertilization with an all-terrain vehicle (ATV) that had a Push Broadcast Spreader (Quad Master ATV Implements) mounted on the back. Five hundred kilograms of 10:1:5 (31) Nitrophoska® fertiliser: 48.33 kg Potassium, 9.69 kg Phosphorus and 96.88 kg Nitrogen, was spread over the field to ensure nutrient supply for the crops in the field. Debris from the field trials of the former season was removed using a rake towed by a tractor as well as applying fertiliser thoroughly into the soil. SAKURA®850 WG (Bayer CropScience), a pre-emergent herbicide was sprayed on the field (118 g/ha: 30 g dissolved in 70 L of water).

The Wintersteiger Rowseed XL planter was used to plant the seeds. Each one of the 180 plots consisted of five rows with 30 cm spacing between rows. There was an additional 15 cm gap on each side of the plot which resulted in a plot width of 1.5 m. The length of the planted rows was 6 m, with an additional 50 cm on each side sprayed with Roundup. The final plot size was, therefore 7.5 m^2 (5m x 1.5 m). The trial was planted on June 4, 2017.

3.4.4. Crop husbandry

The trial was structured for dryland cropping with rainfall during the winter and received no additional irrigation. Top dressing fertiliser, Turbo 31 (Kynoch Fertiliser) (247 kg/ha, supplying 48 kg/ha), was applied.

Integrated pest management (IPM) was performed to prevent the development of pesticides tolerance. Herbicides were sprayed throughout the field trial to manage weeds growing between rows. MCPA 400 SL (Greena®) and Axial® (Syngenta®) were used to spray the field at 2 L/ha and 778 ml/ha respectively. Seven hundred millilitres of MCPA 400 SL and 273 ml Axial® were dissolved in 70 L of water and sprayed with a QM0422 – 100 L bloomless sprayer. The insecticide Chlophyrifos was used at 750 ml/ha, while the fungicide DuettTM [BASF South Africa, active ingredient: 125 g/L epoxiconazole (DMI–fungicides) and 125 g/L carbendazim (benzimidazole)] was used at the recommended 1 L/ha.

3.4.5. Field data collection

Ground measurements were taken with three handheld instruments that measured the CCI, stomatal conductance and LAI as well as measurements with a RPAS. The measurement was taken on 90, 100, 110 and 125 days after the seeds have been planted and represent the flowering stages of wheat.

3.4.5.1. Stomatal conductance

Stomatal conductance was measured by using the SC-1 leaf porometer (Decagon Devices Inc., Pullman, WA), which is a portable non-invasive instrument (Decagon.com, 2017).

To calculate the repeatability, 10 measurements were taken on the same entry for the three replicates. Since the repeatability of the instrument was not reliable, and measurements took too long to obtain data for all entries within the required time, one measurement was taken on the flag leaf in each entry. The device was calibrated on the day of the data collection to ensure accurate results. Once a reading was complete the sensor head was opened to equilibrate the sensors as indicated in the manual.

3.4.5.2. Chlorophyll content index

The CCM-200 (Opti-Sciences, Tyngsboro, Massachusetts, USA), which is a hand-held chlorophyll content meter, was used in this study. This non-invasive instrument calculates chlorophyll content index (CCI) established at absorbance measurements of 660 and 940 nm and the company states that the instrument has an accuracy of ±1.0 CCI units (Richardson *et al*., 2002).

Ten measurements were taken on the same entry for the three replicates to calculate the repeatability. This was done to indicate how many measurements should be taken for each entry during the data collection. Three measurements were taken on the flag leaf in each entry. The device was calibrated throughout the data collection.

3.4.5.3. Light intercepted by plants

The AccuPAR LP-80 (Decagon Devices, Pullman, WA) was used to measure Photosynthetically Active Radiation (PAR) and LAI. PAR was measured above and below the canopy of the wheat plant, which was then used to measure LAI.

To calculate the repeatability, 10 measurements were taken on the same entry for the three replicates. During the data collection, five measurements were taken above and below the canopies of the wheat plants for each entry. The averages were calculated by the instrument.

3.4.6. Remote pilot aircraft system phenotyping

3.4.6.1. Remote pilot aircraft system flight details and platform

The RPAS flights of the wheat trial were performed on day 90, 100, 110 and 125 after the seeds have been planted, on the same day as the ground measurements (Table 3.4). The RPAS platform used for collecting the first season's data was a 3D Robotics Y6 hexacopter (3DR, 2020) that was controlled by ArduCopter software (Mission Planner 1.3.49) and carried three different cameras at once (Table 3.5). The first

camera was a Go Pro Hero 4 with an RGB 5.4 mm flat lens and the last was a GoPro Hero 4 fixed with an IRPro: NDVI-7 lens to capture near infrared (NIR). NDVI was calculated using the formula NDVI = [NIR-Red]/[NIR+Red]], where reflection within the Red wavelength was captured by the RGB-fitted camera. The RPAS operated autonomously at approximately 20 m above ground level of the site of the trial flying 2 m/s. The flight was automated by the Android application Tower (version 4.0.0). The camera was triggered at distance intervals to attain 70% side-lap and 70% front-lap.

Table 0.4: First season dates of data collection in the field

Dates of field phenotyping	**Days after plant**
02-09-2017	90
13-09-2017	100
21-09-2017	110
10-10-2017	125

Table 0.5: Camera specifications

Manufacturer and model	**Resolution (Px)**	**Sensor size (mm^2)**	**Weight (kg)**	**Spectral range (µm)**
GoPro Hero 4 fixed with a IRPro: NDVI-7	3840x2160	41 x 59 x 30	0.83	410-590 nm and 710-780 nm
Go Pro Hero 4 with a RGB 5.4 mm flat lense	3840x2160	41 x 59 x 30	0.83	410-590 nm and 710-780 nm

3.4.6.2. Commercial pipeline for photogrammetry analysis

The RGB and NIR images taken by the RPAS along with the flight logs (record of the flight events) of each RPAS flight was processed using Mission Planner to generate georeferenced images. The time offset between the clock of the RPAS and camera was calculated so that the global positioning system (GPS) coordinate could be accurately included to the metadata of each image. The geotagged images were uploaded to the commercial Aeroview software. The software uses the geotagged images and Google Earth to create orthomosaic tagged image file format (TIFF) files and a 3D model of the surveyed site. A polygon was drawn around each entry using the software. A shapefile was generated using the polygons drawn on their exact GPS location.

The QGIS software (2.18.14 with GRASS 7.2.2) was used to analyse the orthomosaic images generated by Aeroview. The polygons were edited to ensure that each polygons of the shapefile are correlating with the orthomosaic TIFF file. In order to measure the NDVI value of the polygons the render type parameter was changed to “Singleband pseudocolor” and the colour paramerer to “RDYLGN”. The Zonal Statistic plugin was used to measure the mean, minimum and maximum NDVI value of each polygon.

3.4.6.3. In-house pipeline for photogrammetry analysis

3.4.6.3.1. Orthomosaic generation

The commercial Agisoft Photoscan package (Agisoft LLC, St. Petersburg, Russia, 191144) was used to create orthomosaic images. The pipeline consisted of five major steps (1) the alignment of photos (2) building the geometry (3) building texture (4) generating orthomosaic TIFF files and (5) exporting the file.

During the photo alignment step, the software identifies the orientation and position of the camera for each photo simultaneously and process a sparse point cloud model. The SIFT algorithm of the software was used to identify matching points of distinguished features all over the photos. The generic option was applied for pair selection.

Photos which did not overlap with the trial was excluded before generating an orthomosaic. The step that builds the geometry was selected to be an arbitrary surface type. The smooth model parameter was selected to build the dense cloud during the reconstruction phase and all other parameters were kept on default.

In order to build a 3D model, the build texture was selected from the workflow menu. During this step the Orthophoto parameter was selected to ensure that the entire object surface is textured in the orthographic projection. The blending mode that is a part of this step selects the way how the pixel values from different photos will be combined. The average parameter was selected as this one uses the weighted average value of all the pixels from an individual while the weight is dependent on the same parameters that are considered for high frequency component in mosaic mode. The generated orthomosaic image was exported as a TIFF file and saved.

3.4.6.3.2. Classification by the use of spectral indices

In order to classify the spectral indices, the RGB and NIR images of the calibration panel was imported along with both the orthomosaic images to Matlab (Matlab R2018b, Mathworks, Natick, MA, USA). The orthomosaic images were rotated, and the trial site selected. The Otsu's multilevel thresholding method (MOM) was executed to create a threshold to create two distinguishable classes: Plant and soil. This was done to create the ideal threshold in order to maximise the weighted sum of between-class variances. The percentage coverage of the plants was determined by the total of pixels above the threshold value, which was calculated by the percentage green, divided by the total pixels per plot multiplied by 100. NDVI was calculated for each plot by using the mean red and NIR values from the orthomosaic images (Otsu, 1979) (Hunag *et al*., 2015) (Figure 3.3)

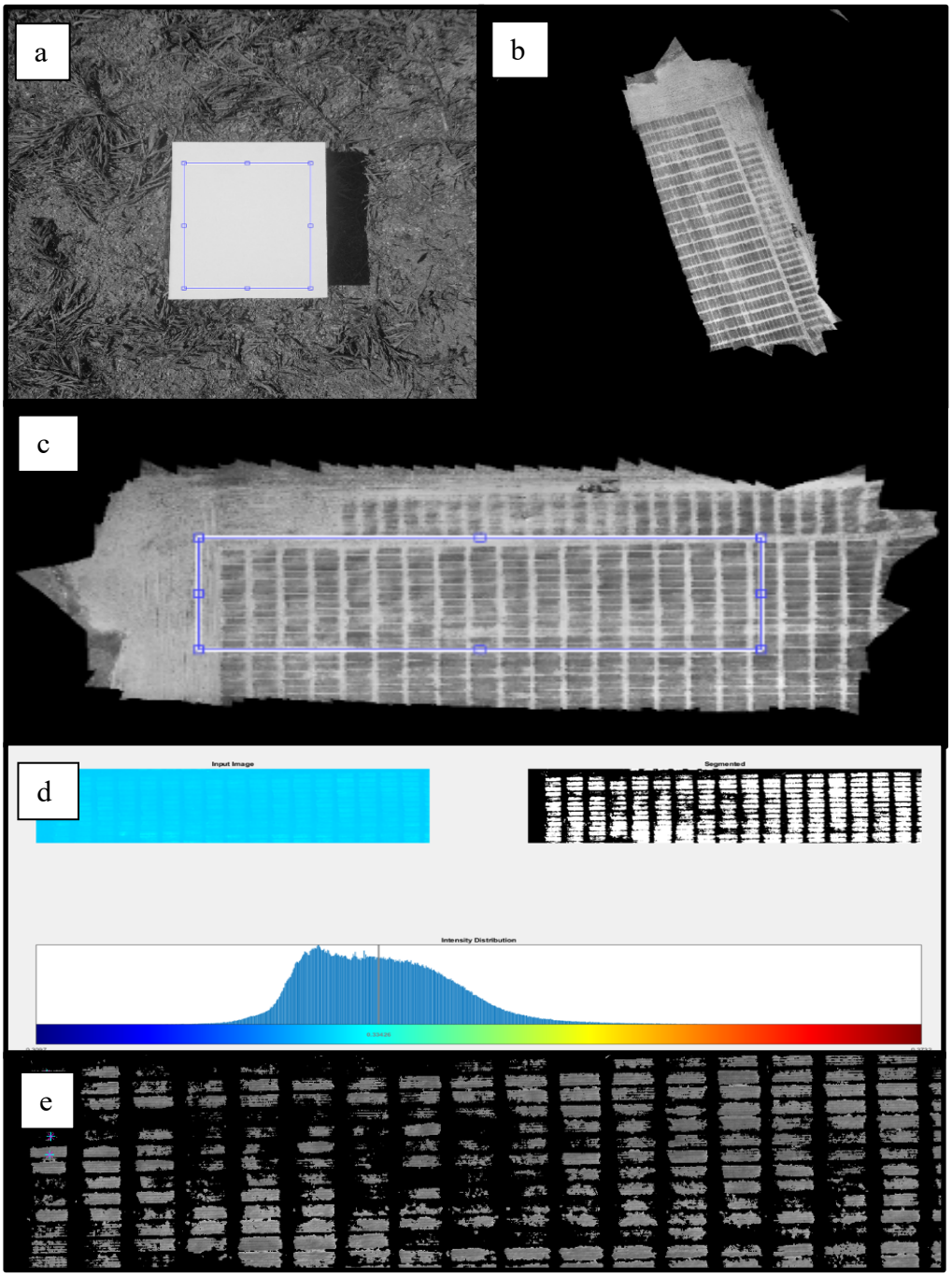


Figure 0.3: The in-house pipeline executed on Matlab a) White reflectance panel pixels selected. b) Orthomosaic of the field trial. c) Othomosaic rotated and the trial site selected. d) Selection of the threshold to distinguish plant and soil pixels. e) Threshold distinguishing the soil and plant pixels from each other.

3.4.7. Harvesting and data analysis

Harvesting of the trial was done on November 23, 2017, almost six months after the seeds were planted. A Wintersteiger Classic plot combine was used to harvest the seeds. The HarvestMaster's Classic GrainGage™ HM800 System fitted to the harvester collected measurements for moisture (%), plot weight (g), test weight (g) and sample sizes larger than 900 g during the harvest. The plot weight and size were used to calculate ton/ha, by calculating the weight per m^2 and multiplying it by 10 000.

After the seeds were harvested and weighed off the field, samples were taken to determine protein (%) and moisture content (%) along with hectolitre (HL) weight by using an Inframatic 9500 NIR Grain Analyzer (Perten, Hägersten, Sweden).

The field data and the data retrieved by the RPAS images were used to create a yield prediction and selection model by using the RStudio (3.5.2) software. The R-squared values were calculated by using multiple linear regression models of the various ground variables fitted on the covariates (RPAS variables). Scatterplot matrices were made to indicate all the tendencies between the ground- and RPAS variables.

3.5. 2018 season's phenotyping of high yielding genotypes

3.5.1. Field trial

The same wheat population as in the first season was planted with a Wintersteiger Rowseed XL planter at Mariendahl experimental farm (33°51'06.25'' S, 18°49'14.14'' E). The layout of the field was randomized and created by using the AgroBase Generation II Version 18.3.1 (Agronomix Software Inc, Winnipeg, Canada) software. Each one of the 180 plots had the same size and layout as the previous season and was planted on June 5, 2018.

3.5.2. Seed treatment

The seeds were treated using the same method as the 2017 season.

3.5.3. Field preparation and planting

The preparation of the field and planting were performed using the same methods as the 2017 season. The same 60 entries were planted in replicates of three.

3.5.4. Crop husbandry

The crop husbandry was performed similar to the 2017 season.

3.5.5. Field data collection

At the experimental site measurements were taken with three handheld instruments that measures the chlorophyll content, stomatal conductance, LAI, leaf temperatures as well as measurements with a RPAS (Figure 3.4). The measurements were taken on 90, 100, 110 and 120 days after the seeds were planted and represented the flowering stages of wheat (Table 3.6)

Table 0.6: Second season dates of data collection in the field

Dates of field phenotyping	**Days after plant**
03/09/2018	90
13/09/2018	100
23/09/2018	110
30/09/2018	125

3.5.5.1. Stomatal conductance

The stomatal conductance was measured in the same way as the previous season, with few changes. Prior to the data collection five measurements were taken on the same leaf of the same plant every hour for eight consecutive hours in order to calculate the repeatability of the device. The device was calibrated before the experiment started each day. During the experiment three measurements were taken on the flag leaf of all the entries of the second replicates.

3.5.5.2. Chlorophyll Content Index

The CCI was measured in the same way as the previous season, with a few changes. Prior to the data collection five measurements were taken on the same leaf of the same plant every hour for eight consecutive hours in order to calculate the repeatability of the device. The device was calibrated before the experiment started each day. During the experiment five measurements were taken on the flag leaf of all the genotypes of the second replicates.

3.5.5.3. Light intercepted by plants

The light interception by the plants were calculated using the same method as the previous season, with a few changes. Prior to the experiment, five measurements were taken above and below the canopy of one entry to calculate the average PAR above and below. This was repeated five times to calculate the repeatability. During the experiment the measurement was repeated two times at each entry of the second replicate.

Prior to the data collection 10 measurements were taken on the same entry for the three replicates to calculate the repeatability. During the experiment, five measurements were taken above and below the canopies of the wheat plants and the averages were given by the device.

3.5.5.4. Leaf temperatures

Leaf temperatures were additionally captured, for the second season. A handheld DT8380 Infrared thermometer was used to measure the flag leaves of the plants. Prior to the experiment five measurements were taken on the same leaf of the same plant every hour for eight consecutive hours in order to calculate the repeatability of the device. During the experiment, three measurements were taken on each genotype of the second replicate.

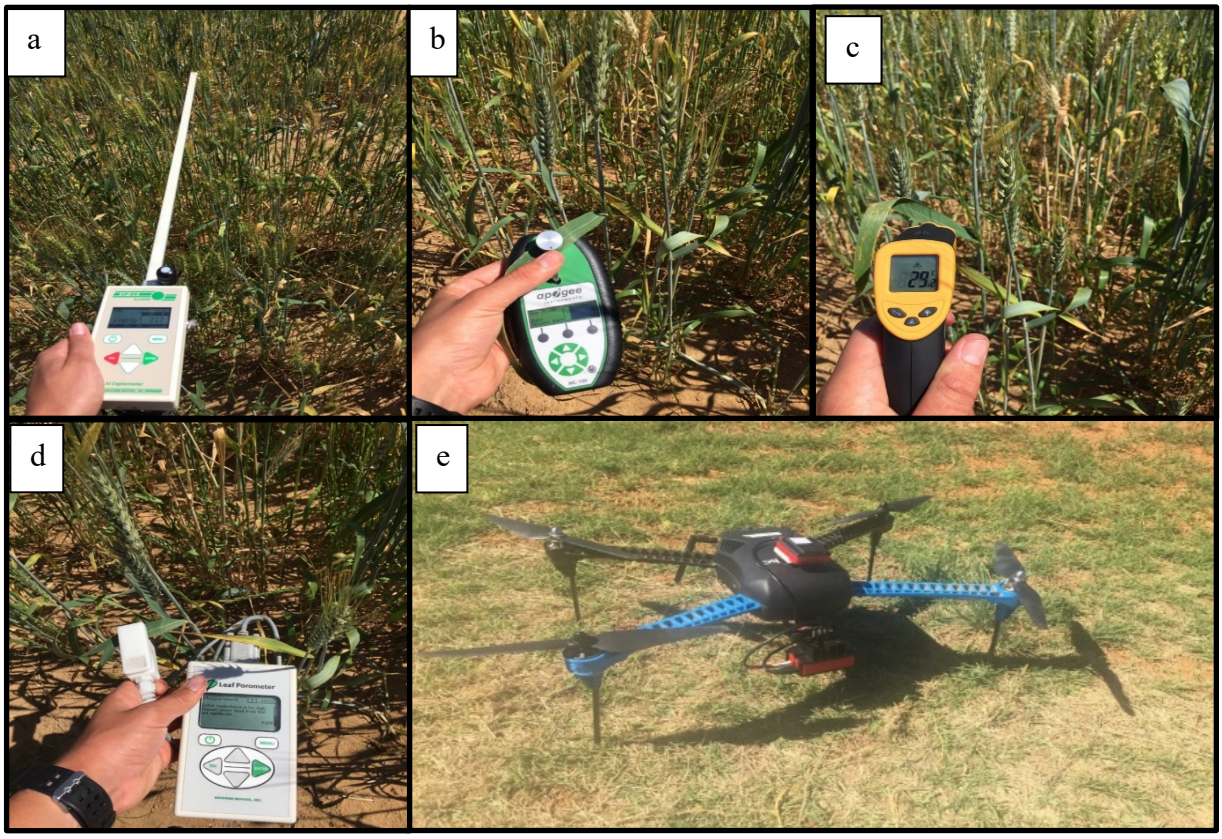


Figure 0.4: Field data collection. a) AccuPAR LP-80 to measure the light interception by plants. b) CCM-200 to calculate CCI. c) DT8380 Infrared thermometer to calculate leaf temperature. d) SC-1 leaf porometer to calculate stomatal conductance. e) Parrot Sequoia multispectral camera mounted to the 3D Robotics IRIS + quadcopter.

3.5.6. Remote pilot aircraft system phenotyping

3.5.6.1. Remote pilot aircraft system flight details and platform

The RPAS flight of the wheat trial was done on day 90, 100, 110 and 120 after the seeds was planted. A single board were placed 1 m above the ground on each corner of the trial that represented ground control points (GCPs). These boards were checked with two matte blocks (25 mm x 25 mm each) and two white blocks (25 mm x 25 mm each) diagonal from each other respectfully. The RPAS used during the second season was a 3D Robotics IRIS + quadcopter and was controlled by the same software as the previous season. The RPAS had a Pixhawk Autopilot hardware (3D Robotics). The three cameras from the previous season were used, and an additional Parrot Sequoia multispectral camera was added to this experiment. This camera has two different sensors which are a multispectral and sunshine sensor in order to analyse

plants by capturing the amount of light they absorb and reflect respectfully. This multispectral sensor has a 16 MPIX RGB camera (4608x3456 pixels) as well as four different 1.2MPIX global shutter single-band cameras (1280x960 pixels) that captures separate bands. The four different bands are Green (550 nm), Red (660 nm), Red Edge (735 nm) and NIR (790 nm). The sunshine sensor has four spectral sensors that are the same as the filters in the cameras. Both the Parrot Sequoia sensors were mounted on the RPAS by using a 3D printed mount (Anon, 2020). The sunshine sensor was mounted on to the aircraft's main centreline facing upwards and the cameras was mounted pointing straight down. The sensor had adequate airflow for cooling during operation as well as a vibration isolator.

Each of the cameras carried on the RPAS at different times. The RPAS operated autonomously at 20 m above ground at 2 m/s. The flight was automated by the Android app namely Tower (version 4.0.0) and the camera was triggered at distance intervals to attain 80% side-lap and 80% front-lap and captured approximately 270 geotagged images for each flight.

3.5.6.2. Orthomosaic generation

Generation of the orthomosaic images from photos captured by the GoPro cameras were done in the same way as the previous season. The photos that were captured with the Parrot Sequoia multispectral camera were stitched with Pix4Dmapper (Version 4.3.31) using the multispectral image analysis settings. An image of the 18% grey calibration card that was captured during each flight was used as a reflectance target during the processing of the images. During the initial processing details step the matching image pairs were set to “Aerial Grid” and it was geometrically verified. For the densification of the point cloud step an optimal multiscale was selected with all the different camera sensors. For the digital surface model (DSM), Orthomosaic and Index details DSM filters was selected to reduce “noise” and the surface smoothing was set to” sharp”. A raster DSM was created with the Triangulation method to merge the tiles. The selected index was NDVI.

3.5.6.3. Classification by the use of spectral indices

The spectral indices obtained from the orthomosaic TIFF file generated by the GoPro images were calculated the same way as the previous season. For the Parrot Sequoia orthomosaics the downstream processing was performed using Python scripts. First the NIR orthomosaic and NDVI images were rotated at the same angle. This was followed by drawing polygons on each of the plots on the rotated NIR orthomosaics and saving the coordinates to a file. A NDVI cut-off was applied to the NDVI images to mask potential non-plant pixels. The percentage coverage and average NDVI for each plot was calculated by using the plot coordinates and masked NDVI images (Figure 3.5).

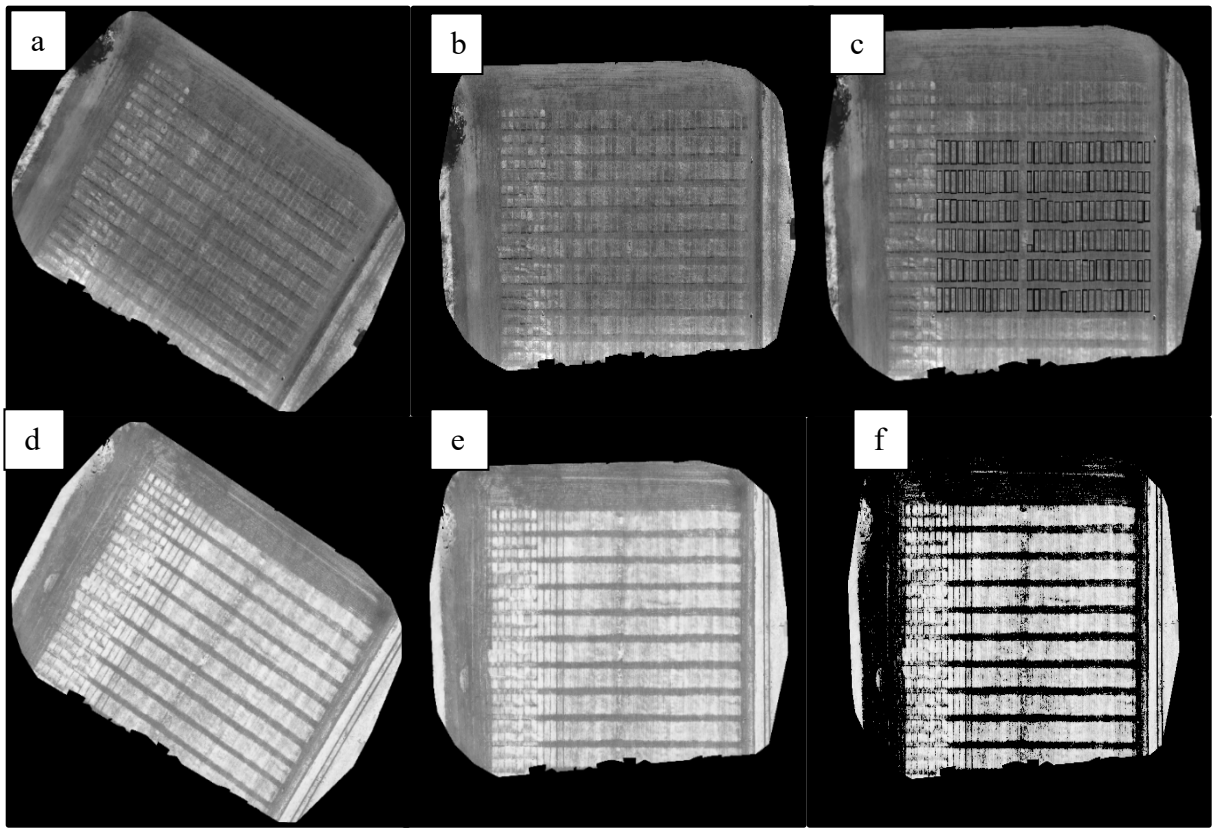


Figure 0.5: The in-house pipeline executed on Python. a) NI orthomosaic generated by Pix4D. b) NI orthomosaic rotated. c) Polygons dragged around each plot to get GPS coordinates. d) NDVI orthomosiac generated by Pix4D. e) NDVI orthomosaic rotated the same angle as the NI orthomosaic. f) NDVI of each pot calculated by using the coordinates of the NI orthomosaic polygons.

3.5.7. Phenotyping high yielding genotypes from the field

Phenotypic data was gathered by pulling plants from the field and analysing them to identify possible high yielding traits (Table 3.7). The height of the plots were measured before the plants were pulled. Three plants were pulled randomly from each of the 60 genotypes in all three replicates to achieve a plot average to measure each of the plot's overall performance. Each tiller of a plant was measured separately by hand, machines and a mobile application, where a scale or ruler was used when required. The Data Count 25-S Plus (DATA Detection Technologies) was used to count seeds, weighing them and calculating TKW and the IM 9500 NIR grain analyser (Perten) was used to measure weight, wet gluten,, protein dry, protein fixed weight basis and moisture (Figure 3.6).

The width, length and area of each individual grain was measured using the SeedCounter Android application. The seeds of each spike were placed on an A4 white sheet with spacing between them to ensure that the application does not identify multiple seeds as one. The images were taken approximately 250 mm above the sheet with a Samsung Galaxy Note4 OS version (Android) 6.0.1 using the application (Figure 3.6). The data was saved and exported as a tsv file on the internal memory of the mobile device. The exported file was opened in Notes and organised by importing it into Microsoft excel (2018).

Table 0.7: How each yield-related trait was measured

Trait	How it was measured	Unit
Plot height	The length of the plot between the base of the tillers and the top of the spike, excluding the awns.	mm
Days to heading	The total days from the plant date till half of the plot have developed a complete head.	Numeric
Plant height	The length between the base of the tiller and the top of the spike, excluding the awns.	mm
Plant weight	Scale	g
Tiller number	Each of the tillers of the plant was counted.	Numeric
Spike length	Ruler (excluding awns).	mm
Spikelet number	Each spikelet was counted for a spike.	Numeric
Grain number per spike	The total grains obtained from each spike.	Numeric
Flowers	The number of florets per tiller	Numeric
Flower fertility	Dividing grains/spike with the number of florets and converting it to percentage.	%
Plot weight	Data Count 25-S Plus.	G
Thousand kernel weight	Data Count 25-S Plus.	G
Harvest index	Calculated by dividing the weight of the grains/plant by the weight of the entire plant including the grains.	%
Grain width, length and area.	SeedCounter application.	mm/mm^2
Hectolitre	Inframatic 9500 NIR grain analyze	%
Moisture	Inframatic 9500 NIR grain analyze	%
Protein	Inframatic 9500 NIR grain analyze	%
Protein dry	Inframatic 9500 NIR grain analyze	%
Wet gluten	Inframatic 9500 NIR grain analyze	%

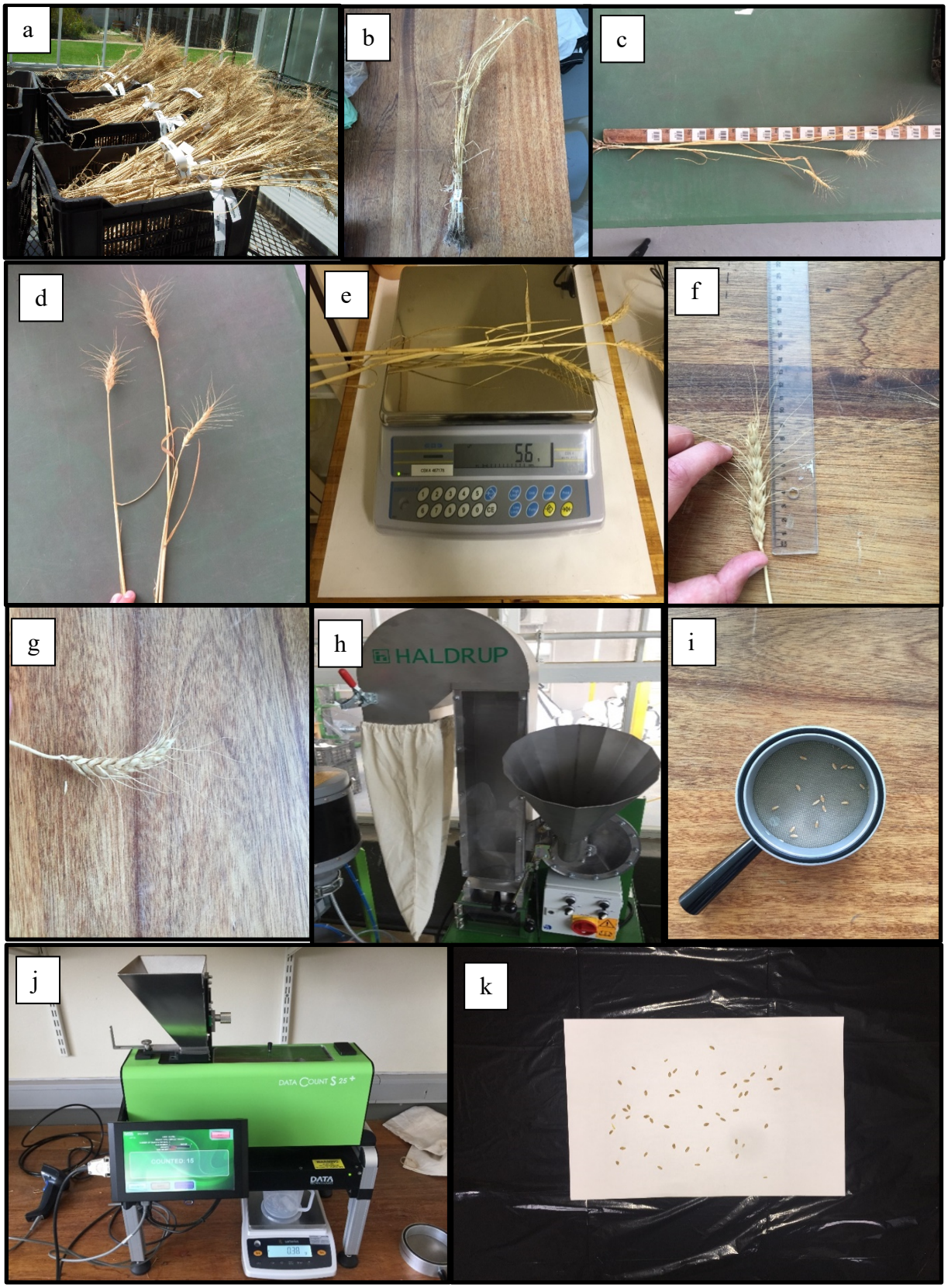


Figure 0.6: Phenotyping high yielding genotypes from the field. a) Plants pulled by hand from field. b) Three plants per plot ready for measurements. c) Length of plant measured. d) Each tiller counted. e) Plant weighed. f) Spike length measured. g) Each spikelet count per spike. h) Tillers threshed to collect seeds. i) Seeds from one tiller. j) Seeds counted, weighed and TKW calculated. k) Seeds on white A4 sheet with a black background for the SeedCounter application.

3.5.8. Harvesting and data analysis

Harvesting of the trial was done on November 26, 2018, almost six months after the seeds were planted. The same equipment was used as in the previous season to harvest and to collect data.

The phenotypic data of the 60 entries with their three replicates each were assembled into an input file and analysed by the Agrobase Generation II version 34.4.18 (Agronomix Software, Winnipeg, Canada) software. Since the measurements were taken for each separate tiller per plant, the values were averaged per plant, followed by averaged per genotype for all three replicates. This was done to create a valid input file that was comprised of averaged measurements for each of the distinctive traits per genotype per replicate. The software was used to test for normality and homoscedasticity of the genotypes. The software was used to do a randomized complete block design (RCBD) ANOVA analysis as it is a statistical analysis which is common in field research. A Nearest Neighbour Analysis (NNA) was also done with the software since this statistical analysis takes spatial variability into account. Outliers were determined and removed by multiplying the interquartile range (IQR) by 1.5 and subtracting the calculated value from the first IQR and adding the calculated value by the third IQR. If a data point did not fall between these two values they were marked as outliers and removed (Taylor, 2019). These outliers were removed since they were most likely caused by human or experimental error and the descriptive statistics are very sensitive to outliers. The field data and the data retrieved by the RPAS images were used to see if there is any correlations between the RPAS data and ground measurements.

Chapter 4: Results and discussion

4.1 Marker-assisted selection

Prior to the MAS, 19 out of the 30 genotypes were selected based on them displaying one or more traits that included plant height TKW, harvest index, grain number per m^2, spikes per m^2, grain number per spike and if the plant height was optimal during the field trial in Obregon Mexico in 2017 (Table 4.2). Three additional CIMMYT cultivars; “Atilla”, “Roelfs” and “Bluebird” were selected to undergo MAS. The selected genotypes were screened by using the standard panel of genetic markers at the SU-PBL (Wessels and Botes, 2014). These markers are used to screen for rust resistance genes (*Sr2, Sr24/Lr24, Sr26, Sr31, Lr19, Lr37/Sr38/Yr17, Lr34/Yr18/Pm38*), the semi-dwarfing gene markers (*Rht-D1b* and *Rht-B1b*) together with baking quality markers (*Glu-Dy10*, Glu-*Dy12* and Glu-*Dx5*). The additional yield-related markers *Ppd-D1*, *TaGS-D1*, *TaGS-3A* and *TaGWS-6B* that were added to the panel in 2017 were also screened for (Rhoda, 2018).

4.1.1. Marker-assisted selection for the rust resistance genes

The stem and leaf rust resistant genes, *Sr2* and *Lr34*, were present in none of the genotypes, which is concerning since these genes are favourable in breeding programs (Figure 4.1). The three resistance genes, *Lr34*, *Lr19* and *Sr26* were also not present in any of the genotypes. The absence of *Sr26* was desired since it is known to have a negative correlation with grain yield and wheat lines that holds these genes are not favourable to include in a breeding program. The *Sr24* and *Lr24* genes occurred in 5% of the wheat. *Sr31* and *Lr37* had a gene frequency of 14% and 18%, respectively. The low frequency of *S31* and *Lr19* in the material is of no concern since both these genes have been overcome by virulent rust pathotypes and are no longer desired in novel germplasm.

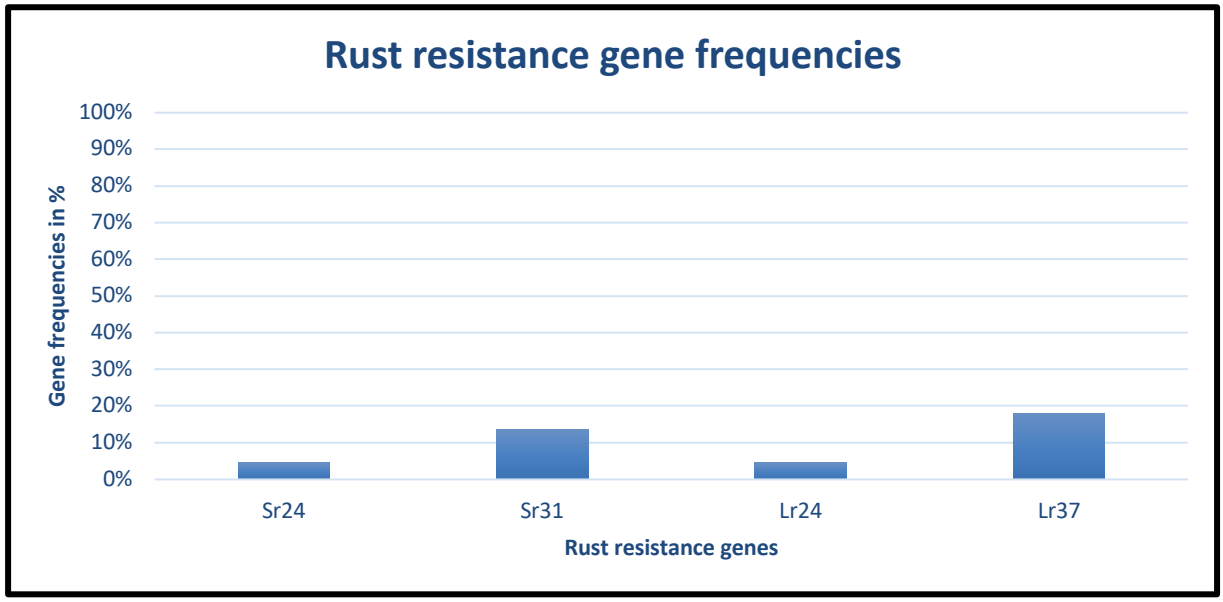


Figure 0.1: Gene frequencies of rust resistance genes.

4.1.2. Marker-assisted selection for the semi-dwarfing and baking quality genes

The dwarfing genes are important to have in any wheat pre-breeding panel of markers, however, only one should be present in the genome. The presence of one of these semi-dwarf genes leads to shorter plants that will prevent lodging, while if both of these semi-dwarf genes were present, the plants would be too short. The marker for the *Rht-B1b* gene were visible in all of the material (Figure 4.2). This is the favoured marker. None of the germplasm had the *Rht-D1b* gene. The *Glu-Dx5* and *Dy10* genes were present in 72.73 % of the genotypes, while the *Dy12* gene was visible in 27.3%

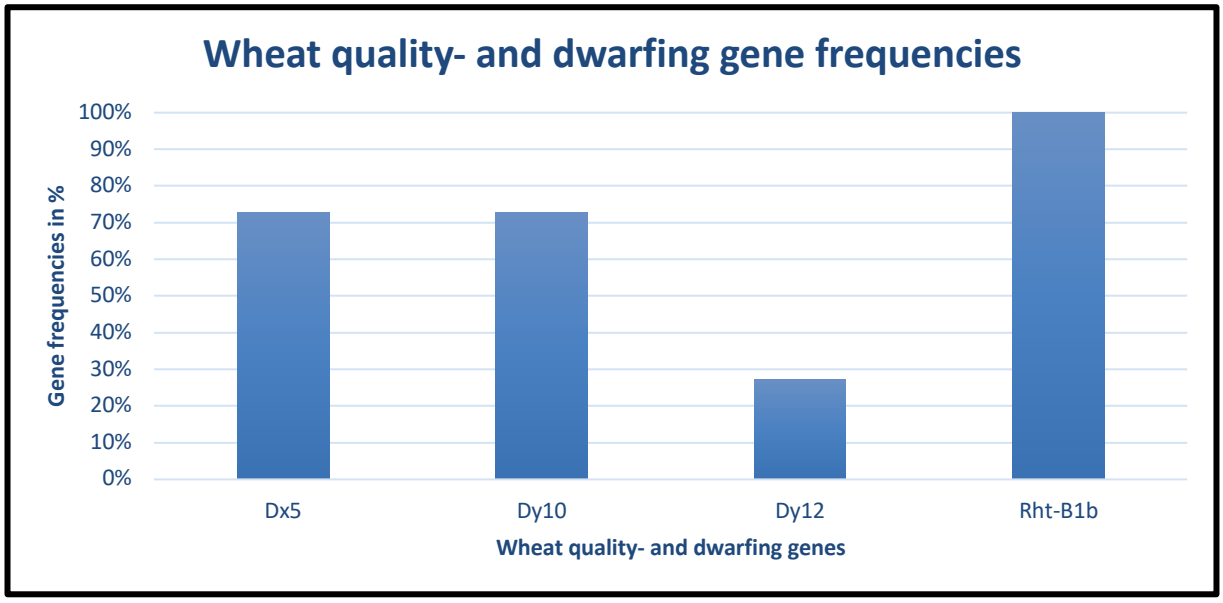


Figure 0.2: Gene frequencies of wheat quality- and dwarfing genes.

4.1.3. Marker-assisted selection for the high yielding traits genes

The 44 samples were screened for genes, which are correlated with traits that can increase grain yield, that were included in the panel of markers in 2017. The *Ppd-D1* gene is correlated with early flowering, which consequently leads to less days to heading. The gene is also linked to the wheat being photoperoid insensitive. The *TaGS-D1*, *TaGW2-6B* and *TaGS5-3A* are associated with the plant presenting the TKW trait, furthermore the *TaGS-D1* gene is also associated with increased grain length trait and *TaGW2-6B* along with *TaGS5-3A* are connected to grain width trait. The genotypes were screened, and the gene frequencies were: 100% for *Ppd-D1*, 66% for *TaGS-D1,* 18% for *TaGW2-6B* and 9% for *TaGS5-3A* (Figure 4.3). The presence of the *Ppd-D1* gene in all of the markers was ideal, since it indicated that all of the genotypes possessed the earlier flowering trait. The high percentage of *TaGS-D1* were in two thirds of the genotypes and was a good indication of which genotypes could be selected during the screening. *TaGW2-6B* and *TaGS5-3A* had low gene frequencies, which was discouraging as they correlate with low grain width.

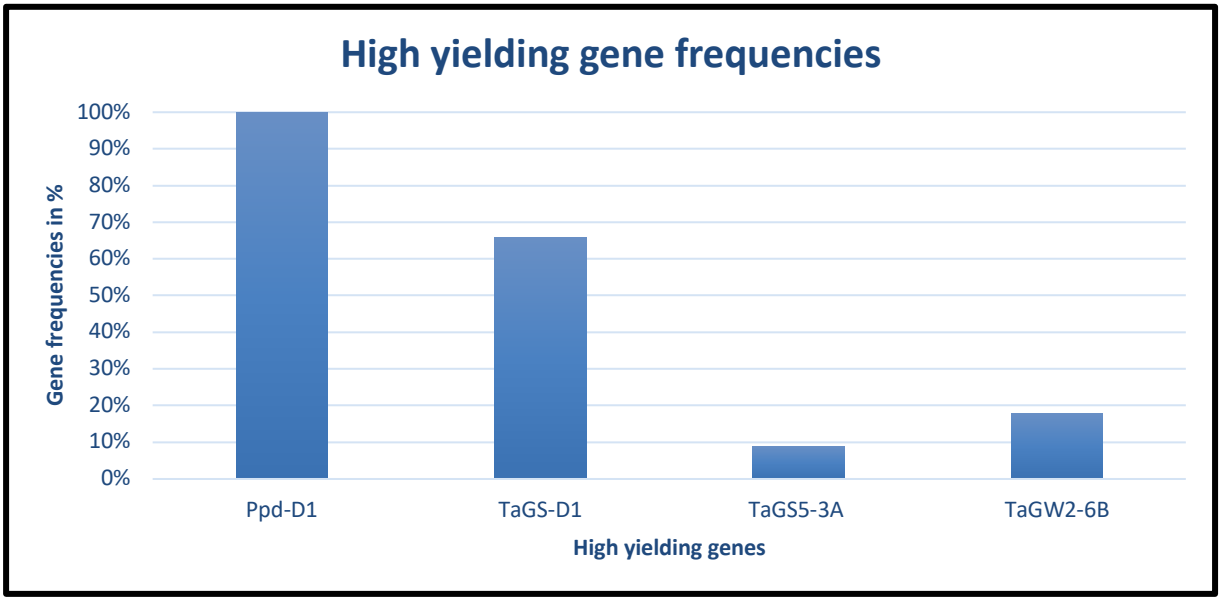


Figure 0.3 Gene frequencies of high-yielding genes.

4.1.4. Genetic diversity assessment

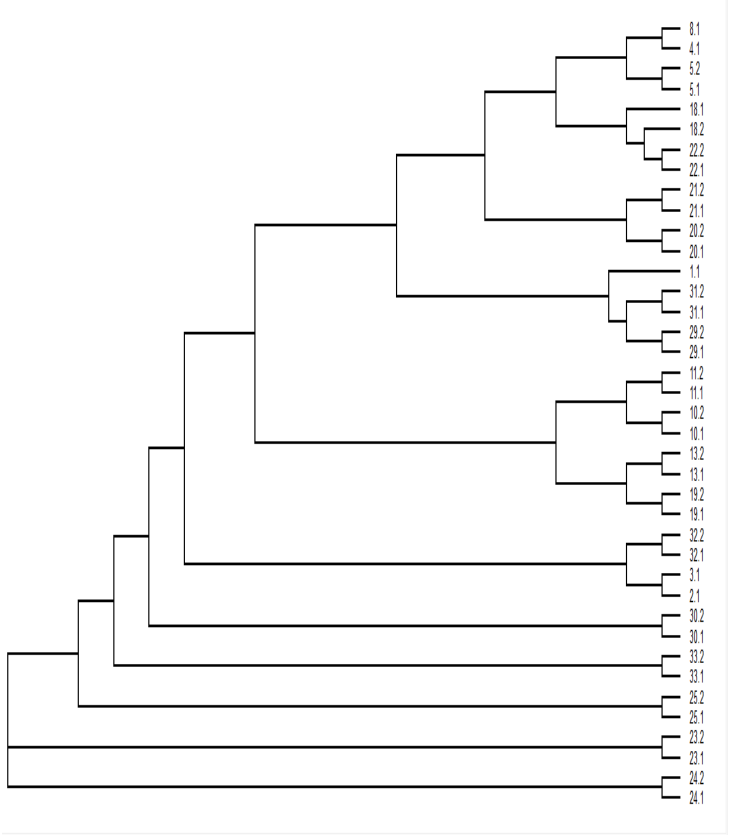
The genetic diversity assessment identified 40 different alleles over the six molecular markers. The polymorphic information content (PIC) values are used to determine how usefull a molecular marker is for linkage studies. The PIC values of the molecular markers ranged between 0.5284 and 0.9329. A PIC value higher than 0.7 is considered highly informartive. Three of the markers (Xgwm190-5D, Xgwm539-2D and Xwmc177-2A) exceeded this value and were the most informative microsatellite markers used in this study (Table 4.1)

Table 0.1: Genetic diversity assessment number of alleles, their range and PIC

Molecular Marker	Nr of alleles	Allele range (bp)	PIC
Xgwm190-5D	18	202-219	0.9166
Xgwm437-7D	3	100-141	0.5807
Xgwm539-2D	3	137-170	0.9329
Xwmc11-1A, 3A	4	172-220	0.6587
Xwmc59-1A, 6A	3	184-220	0.5284
Xwmc177-2A	9	203-239	0.8316

*Base pairs (bp), polymorphic information content (PIC)

The reason the microsatellite markers was to show how similar the wheat lines are to each other, in order to select lines that are more diverse. The Neighbour-joining method was used to construct a distance-based tree, which displayed that all the duplicates of the 22 wheat genotypes grouped together and gave the same results (Figure 4.4). The indicated that the results were accurate since the duplicates had the same results. The tree indicated that each of the duplicates clustered together and one genotype was selected out of each cluster based on the rust, quality and yield markers. This was done to select the most diverse genotypes with the best performance and that possessed wanted genes. There were 13 genotypes selected based on MAS and the results from the 2017 season in Mexico that were included into the second MS-MARS cycle of this study (Table 4.2).



Number	Lines
1	SST0166
2	5th WYCYT 8101513
3	5th WYCYT 8101514
4	5th WYCYT 8101524
5	5th WYCYT 8101533
8	5th WYCYT 8101544
10	5th WYCYT 8101628
11	5th WYCYT 8101631
13	5th WYCYT 8101656
18	5th WYCYT 8101676
19	5th WYCYT 8101677
20	5th WYCYT 8101697
21	5th WYCYT 8101700
22	5th WYCYT 8101703
23	5th WYCYT 8101709
24	5th WYCYT 8101711
25	5th WYCYT 8101712
29	5th WYCYT 7129688
30	5th WYCYT 7705584
31	Atilla
32	Roelfs
33	Bluebird

Figure 0.4: Unweighted pair group method with arithmetic mean (UPGMA) phylogenetic tree for doubled haploid plants.

Table 0.2 Performance of the 19 selected 5th Wheat Yield Collaboration Yield Trial combined analysis under two different sowing dates, Obregon Mexico, 2017 along with the molecular screening. The highlighted entries are the ones selected for the MS-MARS cycle two.

		Rust markers							**Yield markers**				**Gluten markers**			**Height markers**		**2017 Performance**					
Sample Nr	**Sample ID**	*Lr34*	*Sr31*	*Lr24/Sr24*	*Lr37*	*Sr26*	*Lr19*	*Sr2*	*Ppd-D1*	*TaGS-D1*	*TAGS5-3A*	*TaGWS-6B*	*Dx5*	*Dy10*	*Dy12*	*RhtD1b*	*RhtB1b*	TKW	Harvest index	Grain number/m2	Spikes/m2	Grain number/Spike	Plant height
8	8101544								X				1	X			X			X	X		X
4	8101524								X	X			X	X			X			X			
5	8101533									X			X	X			X		X				
18	8101676								X	X					X		X		X				
22	8101703								X	X			X	X			X	X					
21	8101700								X						X		X	X					
20	8101697								X	X			X	X	0		X		X	X	X		
1	SST0166			X	X				X		X	X			X		X						
31	Atilla		X						X		X	X	X	X			X						
29	7129688								X	X			X	X			X				X		X
11	8101631								X	X			0	0	X		X		X			X	
10	8101628								X				0	0	X		X	X					
13	8101656								X				0	0	X		X		X		X		
19	8101677								X				X	X			X		X			X	
32	Roelfs								X	X			X	X			X						
3	8101514								X	X			X	X			X		X	X	X		
2	8101513								X	X			X	X			X			X	X		
30	7705584		X						X				X	X			X	X					X
33	Bluebird		X						X	X			X	X			X						
25	8101712				X				X	X			X	X			X			X		X	
23	8101709				X				X	X		X	X	X			X				X		
24	8101711				X				X	X		X	X	X			X		X	X	X	X	X

4.2. Male sterility marker-assisted recurrent selection scheme

The male sterile plants were phenotypically selected by an evaluation of the tillers. The spike of the tiller had a lighter appearance as well as extensive awns and the anthers was shrivelled and smaller due to the pollen being absent in the male sterile plants. The male sterile plants were selected with caution to prevent the selection of male fertile plants that will lead to selfing. The cross-pollination was effectively executed because of the hybrid seeds having a shrivelled and smaller appearance (Marais & Botes, 2009).

4.2.1. Male sterility marker-assisted recurrent selection cycle one (2017)

The first MS-MARS cycle was done in collaboration with R Rhoda (Rhoda, 2018). The cross-pollination for the first cycle of this study used four high-yielding genotypes that were based on phenotypic data along with the SU-PBL 2017 nursery as the male fertile population and the seeds that were harvested in the previous recurrent cycle were the male sterile population. Overall there were 18 sessions of cutting with the number of male sterile and male fertile tillers alternating between 40-140 and 50-140 respectfully. There were 1577 male fertile tillers used to cross-pollinate and 1811 male sterile tillers. The cycle generated 10004 seeds and the average cross-pollinating rate was 23.95% along with an average grain mass of 238.5 g.

4.2.2. Male sterility marker-assisted recurrent selection scheme cycle two (2018)

The cross-pollination for the second cycle of this study used the genotypes that were selected during the molecular screening as male fertile population and the seeds that were harvested in the previous MS-MARS cycle were the male sterile population. Overall there were 10 sessions of cutting with the number of male sterile and male fertile tillers alternating between 15-80 and 23-70 respectfully. There were 438 male fertile tillers used to cross-pollinate and 423 male sterile tillers. The cycle generated 1878 seeds and the average cross-pollinating rate was 14.44% along with an average grain mass of 24.636 g. The reason for the significant lower number of tillers and seeds in the second cycle was a result of planting the male fertile population too late. Most of the male fertile population were not ready to pollinate the male sterile population when the sterile population was ready.

4.2.3. Success of the cross-pollination

The success rate in percentage of the cross pollination for the first and second cycle was 23.95% and 14.44% (Table 4.5-4.6). These numbers were lower than intended since the aim was to achieve a higher cross-pollination percentage. This was most likely due to selecting the male sterile and male fertile tillers at the incorrect reproductive time. Another reason might be that the male fertile tillers were arranged higher than needed above the male sterile tillers. The structure of the floret could have an effect on the success of the cross-pollination. Tillers with narrow glumes had a lower percentage than those with wider glumes. The structure of a male sterile tiller is, therefore, essential for greater pollen reception.

4.2.4. The inheritance of the male sterility in the recurrent population

It is expected that the recurrent base population of the MS-MARS scheme has a 1:1 ratio male sterile:male fertile. After the seeds were harvested from the segregating population the male sterile and male fertile plants were counted and a chi-square analysis was done. This was necessary to calculate the inheritance of the male sterility gene and if the ratio is correct. A chi-square analysis was done for all four benches respectfully along with an altogether analysis (Table 4.3-4.4).

The chi-square analysis for the first cycle indicated that benches two and four had a p-value greater than 0.05. The benches, therefore, displayed a good fit to the ratio. Bench one and three had separate p-values that were less than 0.05. It, therefore, did not display a good fit to the 1:1 ratio. These results indicated that the single dominant gene within a heterozygous situation that controls male sterility within the population could not be confirmed (Table 4.3). For the second cycle, all the benches had a p-value larger than 0.05 therefore, it displayed that it had a good fit for the 1:1 ratio. This indicated that the single dominant gene within a heterozygous situation that controls male sterility within the population could be confirmed (Table 4.4). The divergence of the benches in the first cycle from the 1:1 ratio could be that the recording of the male sterile and male fertile plants was problematic due to human error.

Table 0.3: The inherited male sterility of the first cycle's recurrent population

Bench	Pots	Sterile plants	Fertile plants	X_2	Probability of fit to a 1:1 ratio
1	120	273	160	29.49	<0.001
2	120	254	231	1.091	0.296
3	121	307	222	13.658	<0.001
4	63	117	117	0	1
Overall	424	951	730	29.055	<0.001

Table 0.4: The inherited male sterility of the second cycle's recurrent population

Bench	Pots	Sterile plants	Fertile plants	X_2	Probability of fit to a 1:1 ratio
1	128	132	142	0.36496	0.546
2	128	111	127	1.07563	0.300
3	128	205	181	1.49223	0.222
4	64	129	120	0.3253	0.568
Overall	448	577	570	0.04272	0.836

Table 0.5: Results of the first male sterile marker-assisted recurrent selection cycle

MS-MARS cycle 1									
Threshing dates	**Week**	**Male Fertile**	**Male Sterile**	**Possible Combinations**	**Harvest**	**Amount of sterile plants**	**Seeds per harvest**	**Average cross pollination (%)**	**Total Mass**
	1	144	251	36144					
26-Jul	2	174	191	33234	1	432	938	8.36	32.8
	3	158	231	36498					
10-Aug	4	187	226	42262	2	431	4470	34.8	91.8
	5	176	199	35024					
31-Aug	6	192	226	43392	3	427	251	22.06	49.2
	7	267	186	49662					
	8	189	252	47628					
06-Nov	9	90	49	4410	4	500	4345	30.58	64.7
Total		1577	1811	328254	4	1790	10004	23.95	238.5

Table 0.6: Results of the second first male sterile marker-assisted recurrent selection cycle

MS-MARS cycle 2									
Threshing dates	**Week**	**Male Fertile**	**Male Sterile**	**Possible Combinations**	**Harvest**	**Amount of sterile plants**	**Seeds per harvest**	**Average cross pollination (%)**	**Total Mass**
	1	67	39	2613					
3-Sept	2	130	89	11570	1	119	422	9.15	8.384
	3	61	77	4697					
10-Sept	4	23	34	782	2	103	1127	17.88	12.434
	5	93	114	10602					
17-Sept	6	64	70	4480	3	176	329	16.3	3.818
Total		438	423	185274		398	1878	14.44	24.636

4.3. Phenotypic data collected in first field trial

The field data, along with the data collected by the RPAS and after the seeds have been harvested was used to create a statistical model for yield predictions and selections of the best high-yielding genotypes. The field and RPAS data were collected on 90, 100, 110 and 125 days after the seeds were planted. The R-squared values in table 4.7 on day 90, indicated that there was no strong correlation between the variables (hand-held measurements) and covariables (percentage coverage measured by two GoPro cameras as well as NDVI) and that a there was not a specific amount of days after the seeds have been planted that had a strong correlation with the variables (Figure 4.5-4.8).

Table 0.7: R-squared values of multiple linear regression models of the various ground variables fitted on the covariates (RPAS variables), COVERNIR (Percentage coverage measured by the NI wavelength) and NDVI for the first trial (2017)

Days after plant	Grain weight	LAI	Stomatal conductance	CCI
90	0.547	0.165	0.010	0.173
100	0.162	0.168	0.066	0.312
110	0.1	0.124	0.07	0.341
125	0.083	0.296	0.07	0.233

*Leaf area index (LAI), chlorophyll content index (CCI)

The R-squared values for LAI indicated that the LAI is best to be measured 125 days after seeds have been planted. The R-squared values for stomatal conductance were very low for all four data collection dates, however, 110 and 125 days after the seeds were planted seemed to be the best for the data collection. The scatter plot matrix indicated that the ground measurements didn't correlate with the variables, NDVI and the percentage coverage measured by the RGB- and NI camera (Table 4.7) (Figure 4.5-4.8).

The data of the commercial pipeline was not used since it did not distinguish the wheat from the soil as in our in-house pipeline. This resulted in incorrect percentage coverage and NDVI values for each plot since data from non-wheat pixels were used in the calculation thereof.

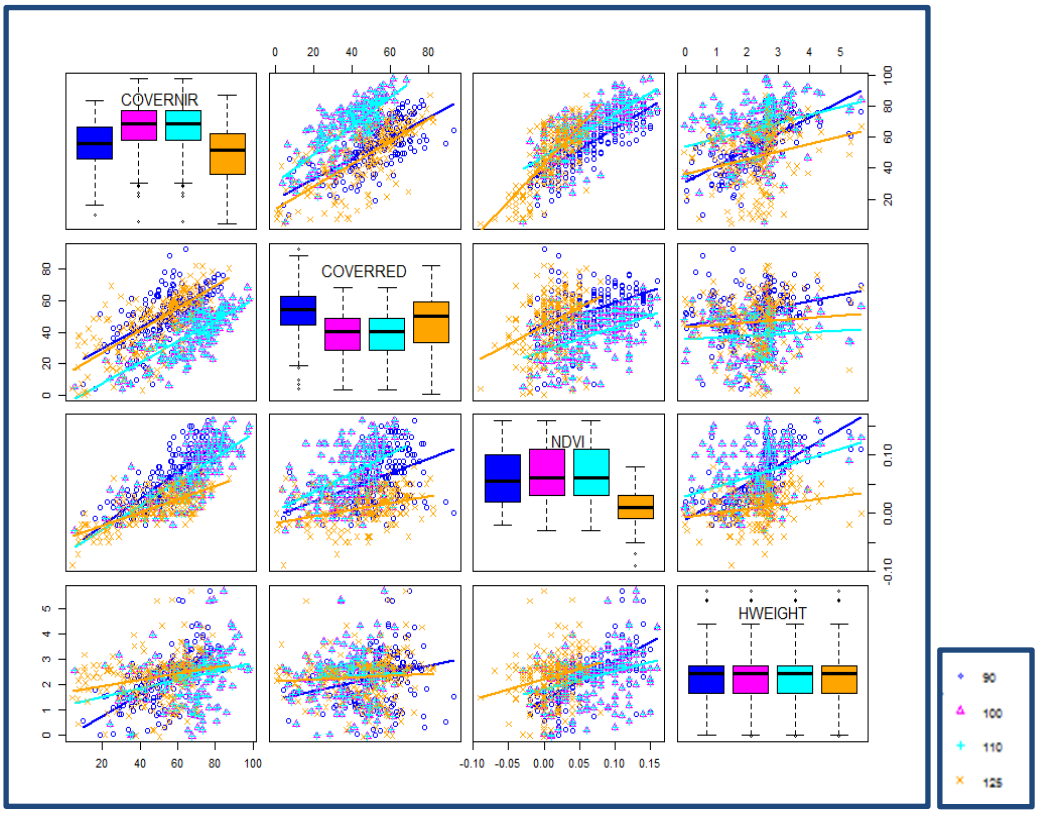


Figure 0.5: Scatter plot matrix for grain weight for first season. The diagonal we have a set of boxplots for the four different time intervals for each variable. The plot gives all pairwise associations between the four variables. These sets are mirror images of one another. The legend on the right is the days after the seeds have been planted.

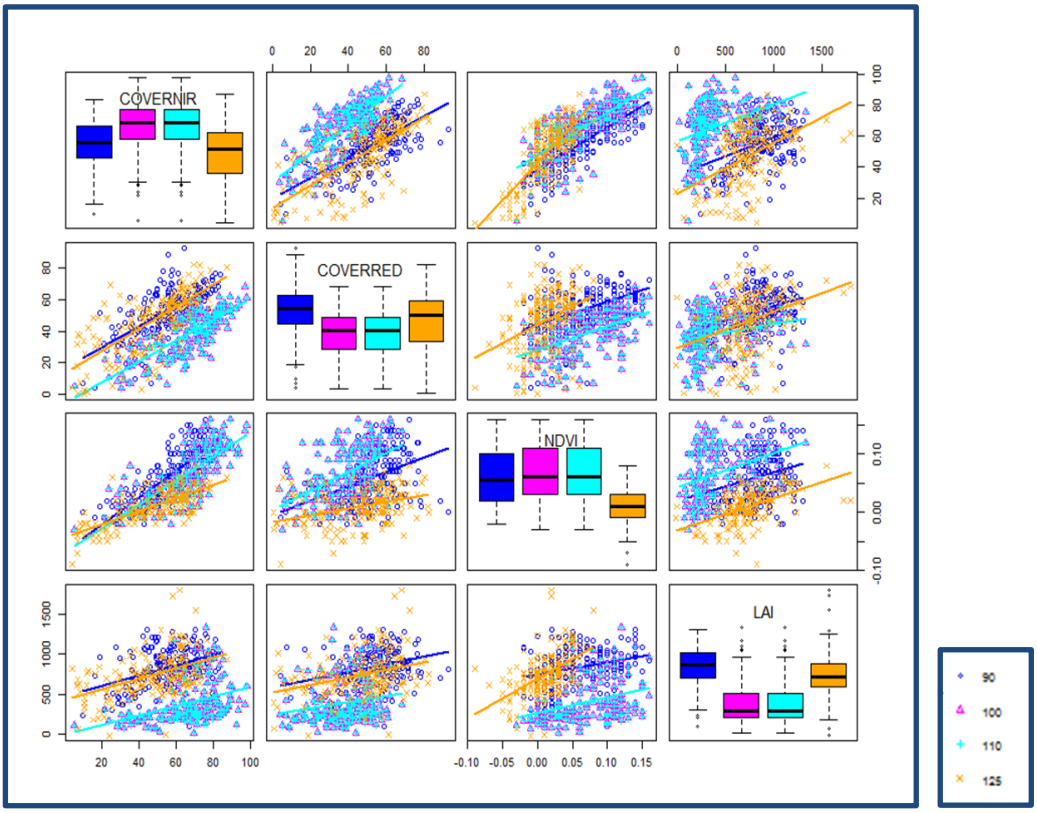


Figure 0.6: Scatter plot matrix for leaf area index for first season. The diagonal we have a set of boxplots for the four different time intervals for each variable. The plot gives all pairwise associations between the four variables. These sets are mirror images of one another. The legend on the right is the days after the seeds have been planted.

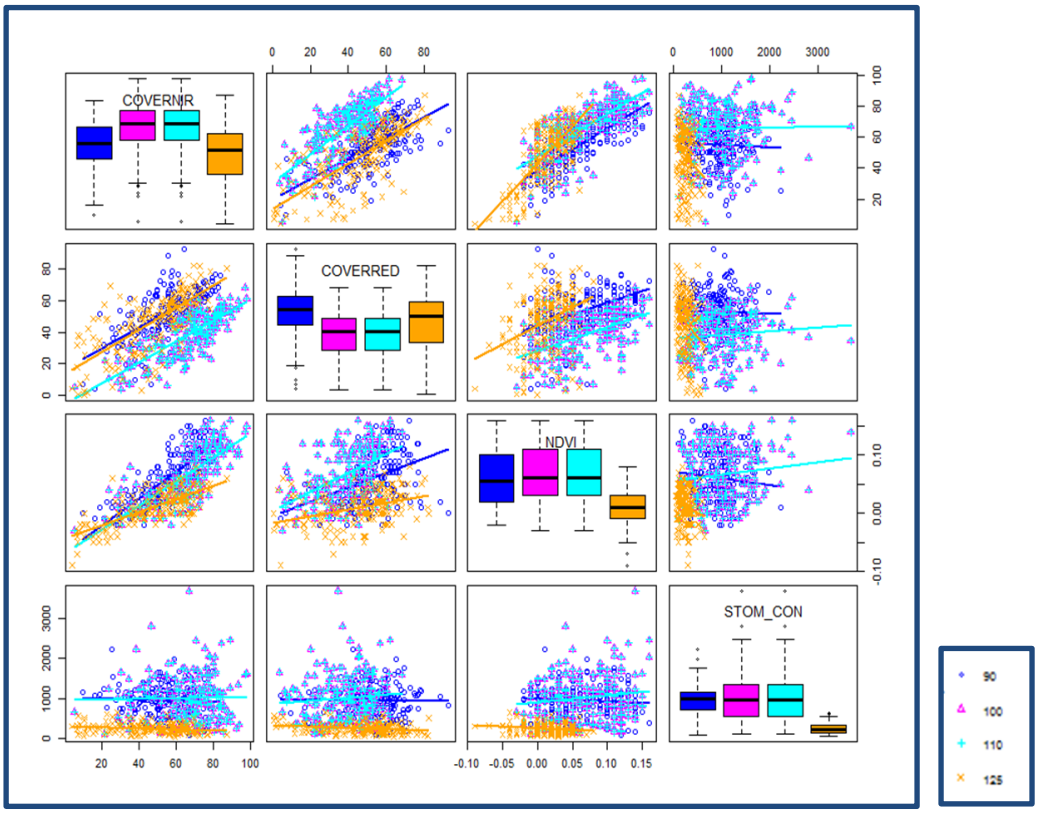


Figure 0.7: Scatter plot matrix for stomatal conductance for first season. The diagonal we have a set of boxplots for the four different time intervals for each variable. The plot gives all pairwise associations between the four variables. These sets are mirror images of one another. The legend on the right is the days after the seeds have been planted.

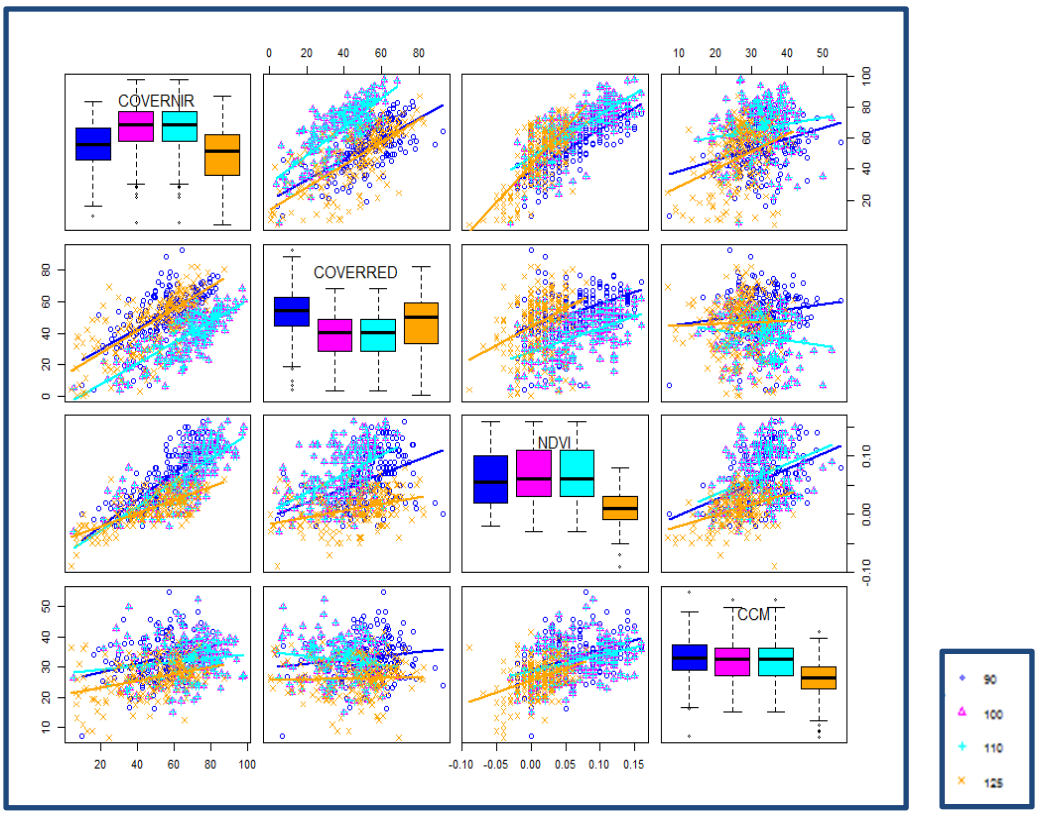


Figure 0.8: Scatter plot matrix for chlorophyll content for first season. The diagonal we have a set of boxplots for the four different time intervals for each variable. The plot gives all pairwise associations between the four variables. These sets are mirror images of one another. The legend on the right is the days after the seeds have been planted.

The plant's hormones have an effect on every part of the development and growth during the life cycle of the plant. Chlorophyll content can be used as an indicator of stress in the plants. This can conclude why the correlation of the chlorophyll content for all four dates were relatively low. A reason for the difference between the replicates could be due to the change of weather during the extended time of measuring for all three ground variables, influencing the results. The sample size should be decreased in order to reduce sampling time, but it should still be statistically acceptable. Another reason could be the severe drought that the trial experienced, which lead to plants not reaching their optimal performance. There was also damage caused by birds, despite the seeds being treated with Anchor® Red that could be seen from the RPAS where the density of the plots were lower than the rest (Figure 4.9-4.10).

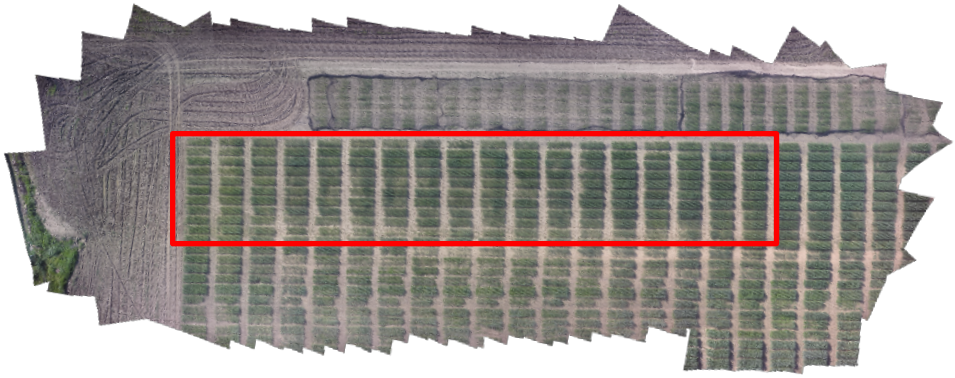


Figure 0.9: Orthomosaic image of the 2017 field trial on 90 days captured by GoPro Hero 4 Silver with red green blue lens. The trial is indicated by the block.

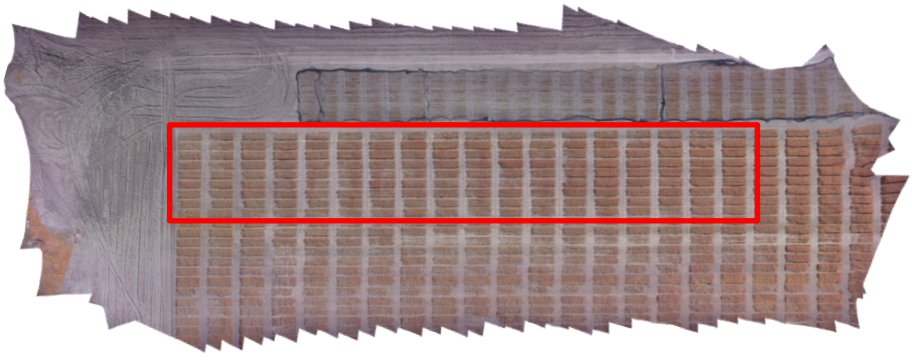


Figure 0.10: Orthomosaic image of the 2017 field trial on 90 days captured by GoPro Hero 4 Silver with near infrared lens. The trial is indicated by block.

4.4. Data collected in the second field trial

During the 2018 season there were an additional variable, temperature, added as well as two new covariables, NDVI and percentage coverage both measured by the Parrot Sequoia multispectral camera. In the first season a very simple camera setup was used with the modified GoPro lenses and in the second season we included the multispectral camera in order to capture more accurate data. R-squared values were measured for both the data that was captured by the GoPro and multispectral camera. The R-squared values of multiple linear regression model indicated that there were no associations between the variables and covariates in this season as these values were very low (Table 4.8). It was expected that there would be associations between the grain weight, LAI, CCI and temperature with all three covariates. For the second season there were also no strong correlations. The data set was a lot smaller than in the previous season, while the conditions of the trial was unfavourable, and this might have been the reason to not seeing any associations. A bigger sample size would have created higher confidence levels. Natural effects could be seen on the aerial images and also have influenced the results and is mainly due to low rainfall during important growth stages in July and August (Figure 4.16) and a lot of damaged plants caused by birds (Figure 4.17-4.18). The percentage coverage for each plot was extremely high since there were grass growing between the rows of the plots. This also influenced the NDVI values, since digitally no distinction could be made between the wheat and other plants. As previously mentioned, in other studies there were correlation between chlorophyll content, stomatal conductance and temperature with NDVI and LAI with percentage coverage (Baresel *et al*., 2017; Bendig *et al*., 2015; Busemeyer *et al*., 2013; Fan *et al*., 2017; Gandhi *et al*., 2015; Haghighattalab *et al*., 2016; Havé *et al*., 2015; Kipp *et al*., 2014; Montesinos-López *et al*., 2017). We did not document it in this study and it is possibly due to natural effects.

Table 0.8: R-Squared values of multiple linear regression models of the various ground variables fitted on the covariates for the second trial (2018)

Ground variable / Time	Grain weight		LAI		Stomatal conductance		CCI		Temperature	
	GP	MS	GP	MS	GP	MS	GP	MS	GP	MS
90	0.015	0.039	0.01	0.04	0.016	0.177	0.017	0.011	0.023	0.025
100	0.059	0.118	0.076	0.019	0.017	0.055	0.014	0.033	0.007	0.023
110	0.029	0.024	0.072	0.031	0.001	0.044	0.064	0.02	0.058	0.035
120	0.06	0.035	0.002	0.006	0.025	0.055	0.017	0.023	0.001	0.035

*Leaf area index (LAI), chlorophyll content index (CCI), GoPro data (GP), Multispectral camera data (MS).

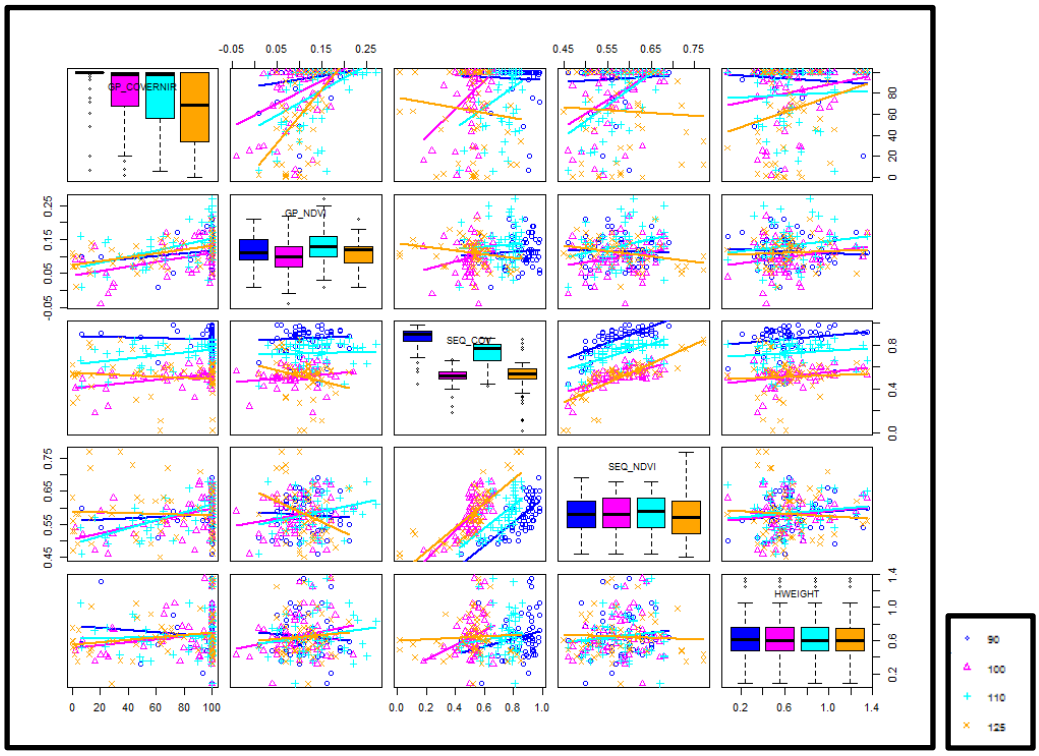


Figure 0.11: Scatter plot matrix for grain weight for 2018 season. The legend on the right is the days after the seeds have been planted.

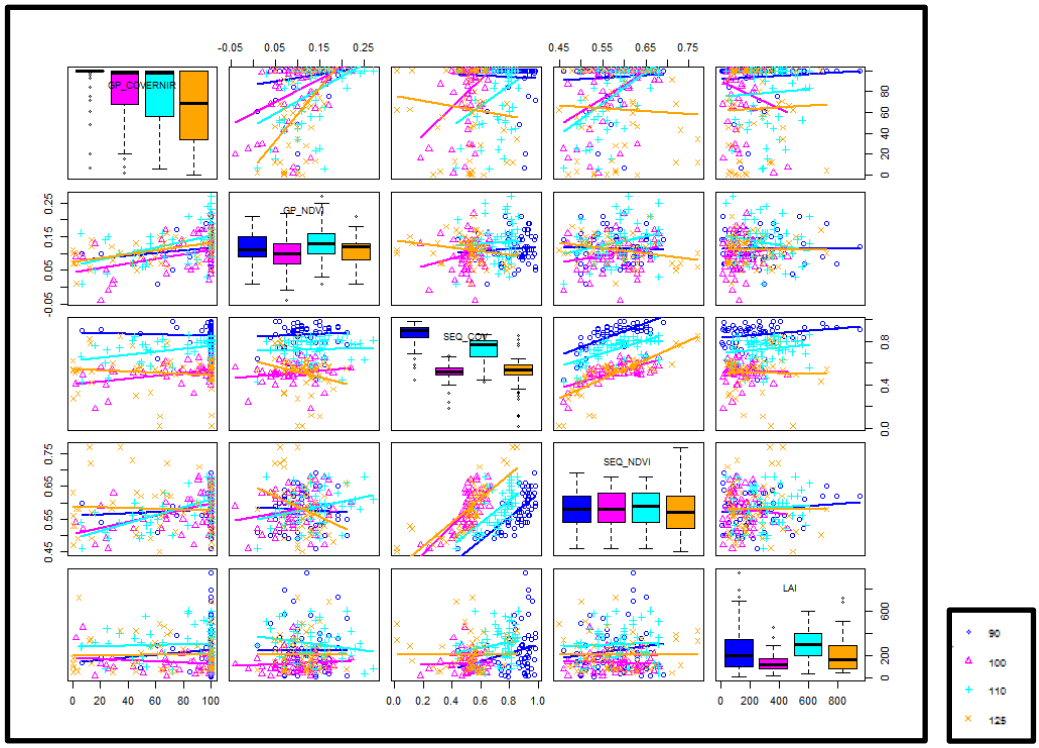


Figure 0.12: Scatter plot matrix for leaf area index for 2018 season. The legend on the right is the days after the seeds have been planted.

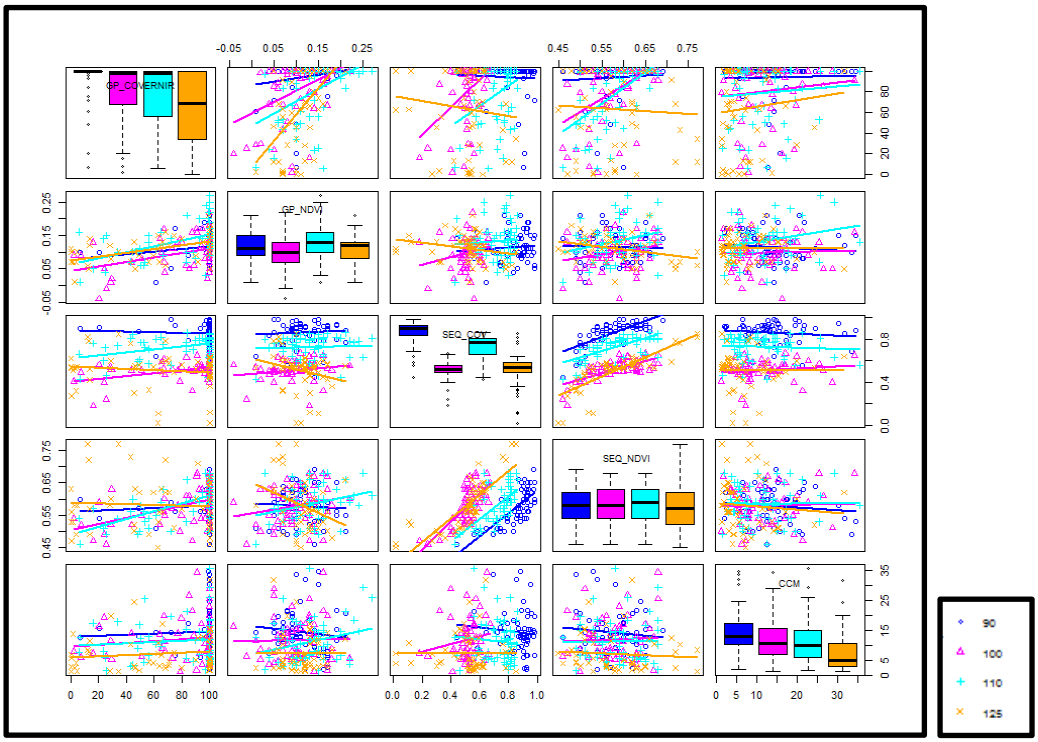


Figure 0.13: Scatter plot matrix for chlorophyll content for 2018 season. The legend on the right is the days after the seeds have been planted.

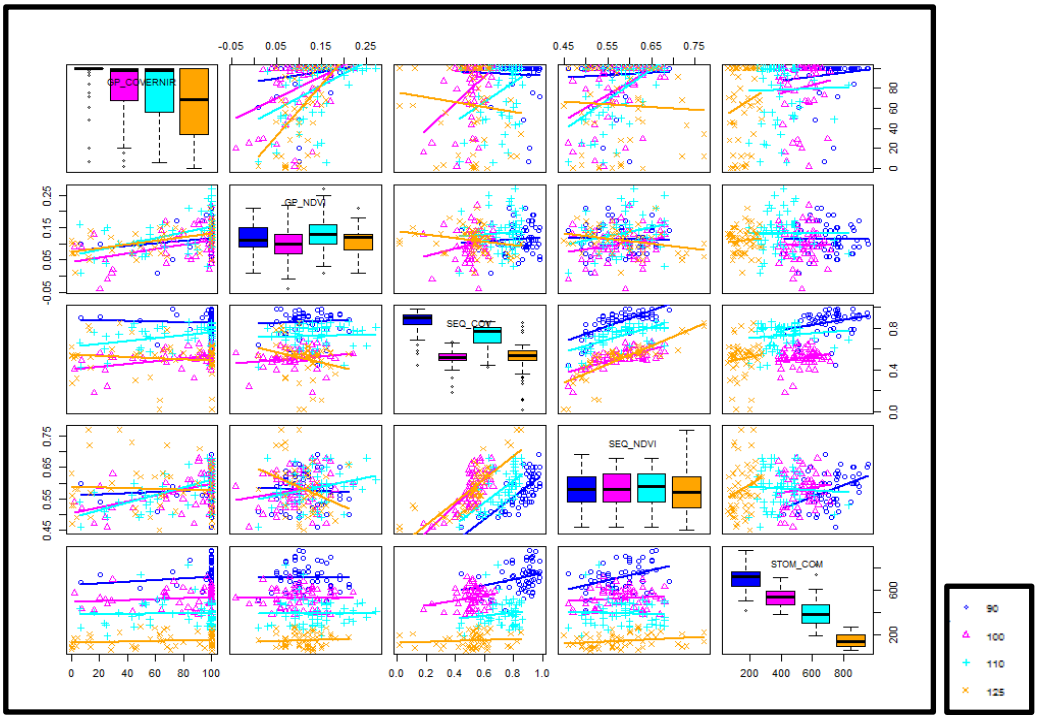


Figure 0.14: Scatter plot matrix for stomatal conductance for 2018 season. The legend on the right is the days after the seeds have been planted.

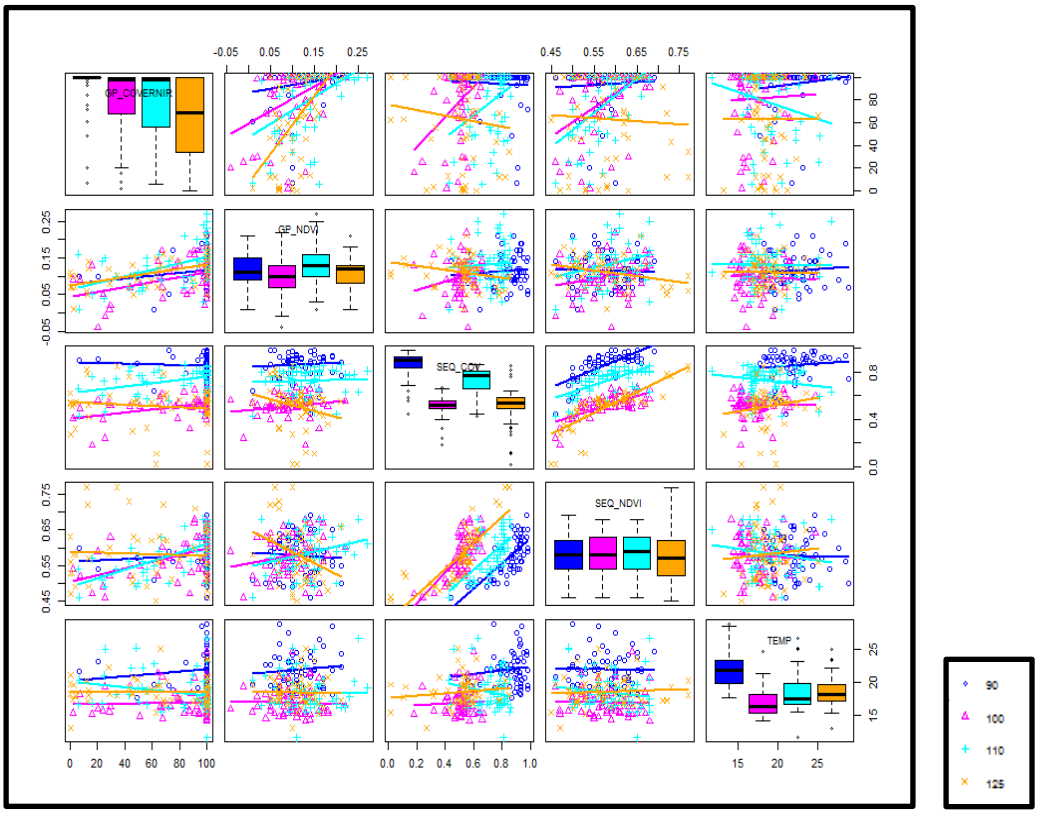


Figure 0.15: Scatter plot matrix for temperature for 2018 season. The legend on the right is the days after the seeds have been planted.

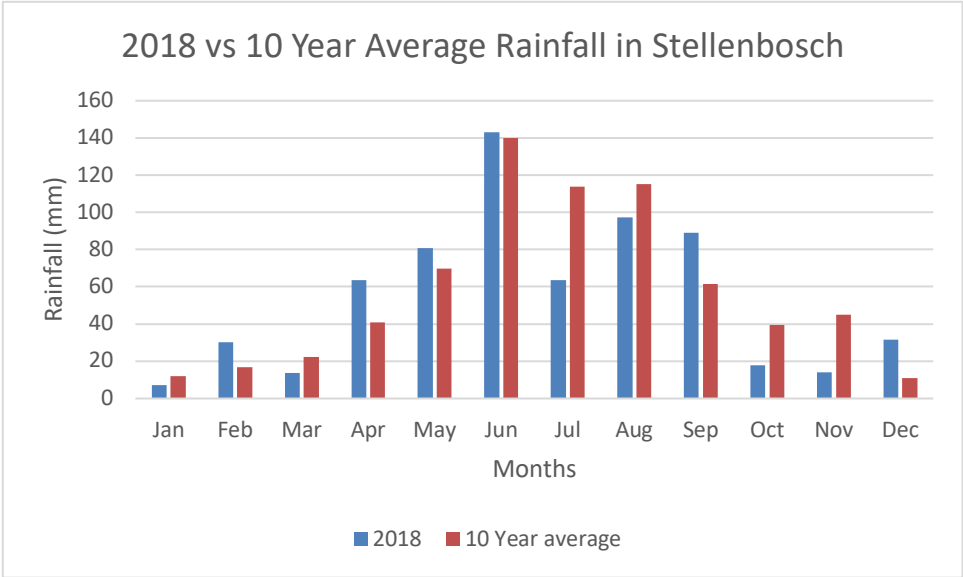


Figure 0.16: 2018 rainfall in Stellenbosch compared to the ten year average for every month

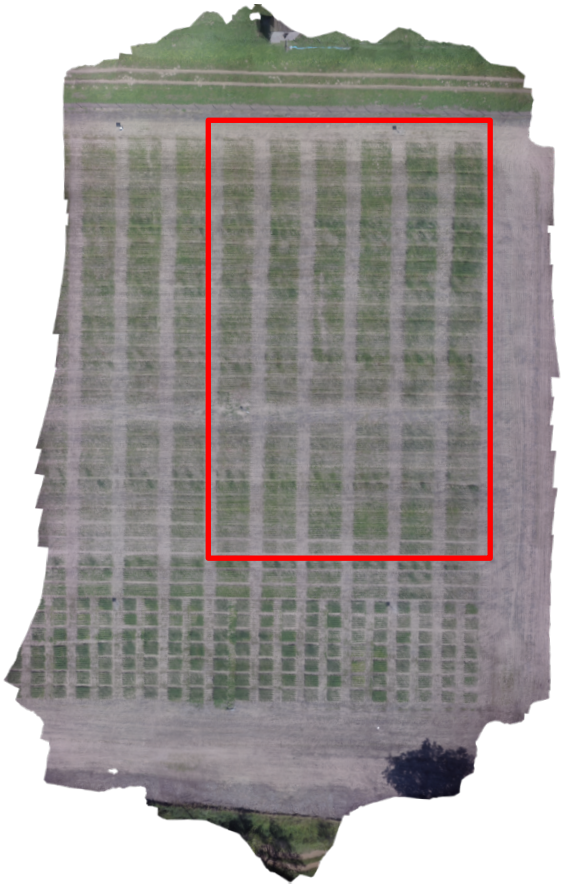


Figure 0.17: Orthomosaic image of the 2018 field trial on 90 days captured by GoPro Hero 4 Silver with red green blue lens. The trial is indicated by the block.

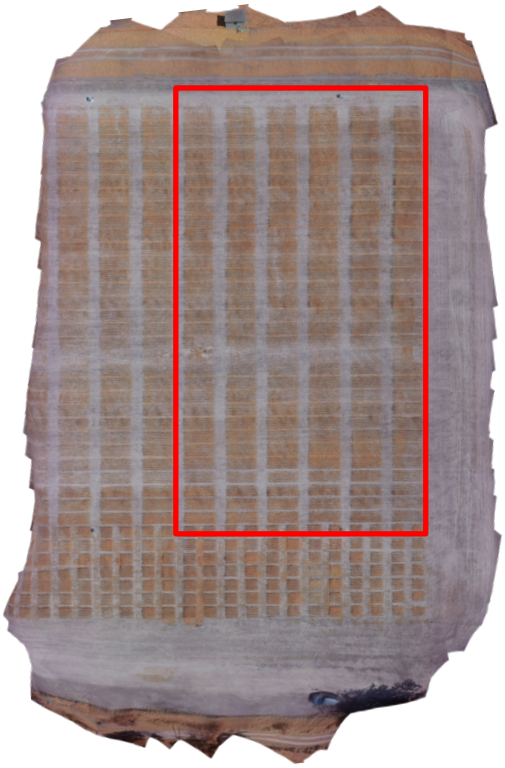


Figure 0.18: Orthomosaic image of the 2018 field trial on 90 days captured by GoPro Hero 4 Silver with near infrared lens. The trial is indicated by the block.

4.5. Phenotyping high yielding genotypes from the field

The Agrobase Generation II version 34.4.18 (Agronomix Software, Winnipeg, Canada) software was used to do a RCBD ANOVA as well as a NNA on the 60 high-yielding genotypes in order to analyse the phenotypic data. Results from the NNA differed from that of the RCBD, which indicated that there were trends observed in the field and it was adjusted for. The data that was generated by the RCBD for the yield- and quality-related traits are presented in Table 4.9. The data generated by the NNA were presented in Table 4.10.

The software produced numerous values, but our main focus was the mean value, coefficient of variation (CV), coefficient of determination (R^2) and broad-sense heritability (H^2) values. The CV is a statistical measurement of the phenotypic dispersion of the different data points around the mean of the dataset. It is calculated by dividing the standard deviation by the mean value times 100. The value is useful to compare the degree of variation between two data series points, even if the means are significantly different from each other (Castagliola *et al*., 2012). The CV values of the RCBD analysis for each individual trait ranged between 1.12% and 33.26%. The CV-values of the NNA for each individual trait ranged between 0.895% and 30.865%. The high CV-values indicated that there was more variation around the mean. The R^2-value ranged between 0.3776 and 0.6747. High R^2-values specifies that the variance in the experiment was explained well by the RCBD, where the low values specifies otherwise. The heritability is important to look at in plant breeding as it indicates the degree of transmissibility of a trait into the prospective generations (Piepho & Möhring, 2007). The software estimated the heritability using the broad-sense heritability equation, Vg/Vp (Falconer & Mackay, 1996). The variation of the entries was used for the genetic variation and the variation of the block was used for the environmental variation. The H^2-values ranged between 0.030 and 0.479. The high H^2-value will indicate that it is effective to select for that trait. The NNA declares more of the residuals in the ANOVA and it is a more suitable analysis of the data if the relative precision is more than 100%. All of the traits had a relative precision of more than 100% and this indicates that the NNA is the better option.

Table 0.9: Randomized complete block design of the phenotyping from field

Randomized complete block design										
Trait	**Grand mean**	**R-squared**	**CV (%)**	**LSD for entry**	**S.E.D**	**p-value (bloc)**	**p-value (Entry)**	**Heritability**	**T (2-sided)**	**MSE**
Days to heading	103.233	0.3776	4.36	7.1179	3.5944	0.0321	0.3366	0.030	1.9803	19.37966
Flowers	34.469	0.4190	16.03	8.9348	4.5119	0.2102	0.0667	0.115	1.9803	30.53603
Flower fertility (%)	69.500	0.5055	13.06	0.1468	0.0741	0.0002	0.0108	0.180	1.9810	0.00824
Harvest index (%)	10.9	0.4044	33.26	0.0585	0.0295	0.4413	0.1468	0.080	1.9814	0.00131
Plot weight (ton/ha)	0.650	0.5890	26.63	0.2799	0.1413	0.0030	0.0000	0.346	1.9808	0.02995
Hectolitre	73.759	0.7129	3.08	3.6880	1.8567	0.1441	0.0000	0.479	1.9864	5.17082
Plot height (mm)	635.519	0.5400	8.91	91.5875	46.2331	0.6050	0.0001	0.302	1.9810	3206.25608
Moisture	11.514	0.5461	1.12	0.2103	0.1058	0.0011	0.0542	0.134	1.9886	0.01678
Plant height (mm)	598.497	0.4821	9.16	88.6468	44.7570	0.4935	0.0035	0.212	1.9806	3004.78079
Plant weight (g)	4.852	0.5222	28.610	2.2456	1.1335	0.8041	0.0004	0.266	1.9812	1.9270
Protein (%)	12.238	0.5940	11.14	2.2110	1.1131	0.3529	0.0003	0.289	1.9864	1.85851
Protein dry (%)	13.904	0.5936	11.19	2.5232	1.272	0.3914	0.0003	0.289	1.9864	2.42025
Seeds per tiller	24.674	0.4868	23.46	9.3592	4.7254	0.0127	0.0071	0.192	1.9806	33.49375
Seed area (mm^2)	7.385	0.5009	8.53	1.0190	0.5145	0.0000	0.0743	0.111	1.9806	0.39702
Seed length (mm)	4.874	0.6747	3.26	0.2571	0.1298	0.0000	0.0000	0.397	1.9810	0.2526
Seed weight (g)	0.504	0.4198	30.8	0.2514	0.1268	0.3938	0.1156	0.092	1.9820	0.02413
Seed width (mm)	1.961	0.4956	5.23	0.1660	0.0837	0.0029	0.0214	0.160	1.9824	0.01052
Spikelets	11.440	0.4265	15.765	2.8945	1.4614	0.4505	0.0500	0.127	1.9806	3.20348
Spike length (mm)	61.242	0.5900	14.07	13.9352	7.0370	0.0000	0.0007	0.251	1.9803	74.27965
TKW (g)	18.688	0.4413	18.57	5.6173	2.342	0.1164	0.0716	0.114	1.9820	12.04717
Tiller number	3.032	0.4347	25.73	1.2625	0.6369	0.5877	0.0701	0.115	1.9822	0.60853
Wet gluten (%)	27.828	0.5930	12.99	5.8627	2.9515	0.3630	0.0003	0.288	1.9864	13.06672

*Thousand kernel weight (TKW), least significant difference (LSD), standard error of a difference between two means (S.E.D), mean squared error (MSE)

Table 0.10: Nearest neighbour anlysis of the phenotyping from field

Nearest neighbour analysis									
Trait	**Grand mean**	**Relative precision (%)**	**CV (%)**	**LSD for entry**	**S.E.D**	**p-value (bloc)**	**p-value (Entry)**	**Heritability**	**T (2-sided)**
Days to heading	103.233	102.770	4.187	6.9902	3.53	0.1054	0.2170	0.058	1.980
Flowers	34.469	122.385	14.448	8.053	4.07	0.5457	0.0025	0.220	1.980
Flower fertility (%)	69.500	103.018	12.593	0.1415	0.07	0.0001	0.0023	0.223	1.980
Harvest index (%)	10.9	108.762	30.865	0.0542	0.03	0.5990	0.0127	0.174	1.980
Weight (ton/ha)	0.650	106.162	25.401	0.2699	0.13	0.0692	0.0000	0.399	1.980
Hectolitre	73.759	157.080	2.151	2.5656	1.30	0.8428	0.0000	0.703	1.980
Plot height (mm)	635.519	140.590	7.353	75.5673	38.16	0.8654	0.0000	0.456	1.980
Moisture	11.514	111.102	0.895	0.1667	0.08	0.1302	0.0000	0.370	1.980
Plant height (mm)	598.497	122.957	8.157	78.9464	39.869	0.8456	0.0001	0.297	1.980
Protein (%)	12.238	192.233	7,013	1.3880	0.70	0.9698	0.0000	0.439	1.980
Protein dry (%)	13.904	192.860	7.032	1.5813	0.80	0.9740	0.0000	0.441	1.980
Seeds per tiller	24.674	110.172	22.079	8.8094	4.45	0.2304	0.0010	0.2304	1.980
Seed area (mm^2)	7.385	113.048	7.929	0.9469	0.48	0.0058	0.0093	0.183	1.980
Seed length (mm)	4.874	113.763	2.991	0.2357	0.12	0.0192	0.0000	0.460	1.980
Seed weight (g)	0.504	126.355	26.199	0.2137	0.11	0.7633	0.0006	0.254	1.980
Seed width (mm)	1.961	124.288	4.447	0.1410	0.07	0.0240	0.0000	0.386	1.980
Spikelets	11.440	115.235	14.412	2.6660	1.35	0.6960	0.0035	0.211	1.980
Spike length (mm)	61.242	128.247	12.398	12.2782	6.20	0.0131	0.0000	0.359	1.980
TKW (g)	18.688	126.552	15.791	4.7724	2.41	0.4807	0.0007	0.252	1.980
Tiller number	3.032	101.336	24.340	1.1935	0.60	0.3750	0.0050	0.201	1.980
Wet gluten (%)	27.828	193.445	8.150	3.6686	1.85	0.9712	0.0000	0.437	1.980

*Thousand kernel weight (TKW), least significant difference (LSD), standard error of a difference between two means (S.E.D), mean squared error (MSE)

The results for the days to heading trait displayed desirable results with a CV-value of 4.36%, R^2-value of 0.3776 and H^2-value of 0.030 (Table 4.9). The results of the NNA for this trait had a CV-value of 4.187, a relative precision of 102.770% and an H^2-value of 0.058 (Table 4.10). This indicated that the NNA was a better option than the RCBD analysis and the field trends do have an influence. These low CV-values proposed that the trait is reliable and should be used in the breeding programme for achieving higher yield. As previously mentioned, it is desired that the wheat starts heading at in less days. The average days it took the wheat to start heading at the experimental site was 103.233 days, where the best performing entries were 12 (96.182), 15 (96.217), 16 (96.723), 18 (98.254) and 56 (98.990) (Table 4.11). All of the entries in the top five were significantly better than the average and they are good candidates to select for as crossing parents.

Table 0.11: Top five candidates for the days to heading trait

Rank	**HYLD Entry number**	**Mean days to heading**	**CV (%)**
1	12	96.182	3.5
2	15	96.217	5.9
3	16	96.723	5.5
4	18	98.254	0.7
5	56	98.990	6.9

The RCBD results for the flowers per spike presented that it is good to be used in a breeding programme as it had a CV-value of 16.03%, R^2-value of 0.4190 and a H^2-value of 0.115 (Table 4.9). The results of the NNA had a CV-value of 14.448, relative precision of 122.385% and a H^2-vale of 0.220% (Table 4.10). The NNA indicated that the field trends have an influence on this trait and this analysis should be applied. The average of all the entries were 34.469, where the five best candidates were 13 (43.116), 48 (42.390), 15 (42.324), 54 (40.912) and 32 (40.731) (Table 4.12). The top five candidates were significantly higher than the average of all the entries. More flowers per spike is a good trait of high-yielding genotypes and the 15, 54 and 32 entries can be used as good crossing parents due to their high number of flowers per spike and the low CV-values of 7.2%, 9.7% and 8.6% respectfully. The low CV-values indicates that the variation between the replicates were low.

Table 0.12: Top five candidates for the flowers per spike trait

Rank	HYLD Entry number	Mean	CV (%)
1	13	43.116	13.6
2	48	42.390	11.1
3	15	42.324	7.2
4	54	40.912	9.7
5	32	40.731	8.6

The RCBD analysis for fertility of the flowers displayed a CV-value 0f 13.06% and R^2-value of 0.5055 and a H^2-value of 0.180 (Table 4.9). The NNA had a CV-value of 12.593%, relative precision of 103.018% and a H^2-value of 0.223 (Table 4.10). The NNA were the better option as the field trends had an influence on this trait. The average of all the entries were 69.500%, where the top five candidates were 9 (84.020%), 18 (83.974%), 38 (81.250%), 17 (81.226%) and 21 (79.263%) (Table 4.13). The low CV-value and high R^2-value indicates that this is a good trait to select for in breeding programmes, however the H^2-value indicated that the heritability for this trait is low. The best entries to select for is entry 9 and 38 due to their low CV-values of 4.9% and 3.4% respectfully.

Table 0.13: Top five candidates for the flower fertility

Rank	HYLD Entry number	Mean flower fertility (%)	CV (%)
1	9	84.020	4.9
2	18	83.974	5.8
3	38	81.250	3.4
4	17	81.226	7.9
5	21	79.263	9.9

The HI results of the RCBD indicated that the trait is desirable to select for in breeding programmes and had a CV-value of 33.26%, R^2-value of 0.4044 and H^2-value 0.080 (Table 4.9). The NNA for this trait had a CV-value of 30.865%, relative precision of 108.762 % and an H^2-value of 0.174 (Table 4.10). The NNA were a better option than the RCBD as field trend had an influence on this trait. Overall, the HI in this experiment were low with an average of 10.9 % with the top five candidates being 31 (17.1 %), 38 (16.7 %), 32 (16.4 %), 48 (16.2 %) and 24 (15.6 %) (Table 4.14). The top candidates displayed a significant higher HI than the average. The trait should not be used as a selection tool for this experiment due to the high CV-values.

Table 0.14: Top five candidates for the harvest index trait

Rank	HYLD Entry number	Mean HI (%)	CV (%)
1	31	17.1	14.6
2	38	16.7	16.9
3	32	16.4	22.5
4	48	16.2	10.6
5	24	15.6	8.2

*Harvest index (HI)

The results of the RCBD for the plot weight had a CV-value of 26.63%, R^2-value of 0.5890 and an H^2-value of 0.0346 (Table 4.9). The results of the NNA had a CV-value of 25.401%, relative precision of 106.162% and H^2-value of 0.399 (Table 4.10). The NNA were the better option as the field trends had an influence on this trait. This indicates that the CV-value of the trait is high, and it might not be effective to select for it. The average weight per plot of this experiment were very low at 0.650 ton/ha, where the five best candidates were 58 (1.016 ton/ha), 18 (0.957 ton/ha), 15 (0.926 ton/ha), 20 (0.925 ton/ha) and 60 (0.919 ton/ha) (Table 4.15). The top five candidates were significantly higher than the average of the experiment.

Table 0.15: Top five candidates for the plot weight of the entire plot, measured by the harvester

Rank	HYLD Entry number	Mean plot weight (ton/ha)	CV (%)
1	58	1.016	13.9
2	18	0.957	22.7
3	15	0.926	27.1
4	20	0.925	20.9
5	60	0.919	11.0

*Ton per hectare (ton/ha)

The RCBD results for the hectolitre trait had with a CV-value of 3.08%, R^2-value of 0.479 and H^2-value of 0.1441 (Table 4.9). The NNA had a CV-value of 2.151%, relative precision of 157.080% and a H^2-value of 0.703 (Table 4.10). The NNA was a better option than the RCBD analysis. The average of all the entries were 73.759 hL, where the top five candidates were 37 (79.286 hL), 19 (78.080 hL), 23 (77.453 hL), 12 (77.295 hL) and 22 (77.281 hL) (Table 4.16). All of the top five candidates were significantly higher than the average. It would be beneficial to select for these five entries as crossing parents in future experiments since they displayed low CV-values and a high heritability.

Table 0.16: Top five candidates for the hectolitre trait

Rank	HYLD Entry number	Mean hectolitre (hL)	CV (%)
1	37	79.286	1.1
2	19	78.080	2.8
3	23	77.453	1.6
4	12	77.295	1.6
5	22	77.281	1.2

*Hectolitre (hL)

The RCBD results for the plot height in the field trait displayed desirable results with a CV-value of 8.91%, R^2-value of 0.540 and H^2-value of 0.302 (Table 4.9). The NNA was a better option as it had a CV-value of 7.353%, a relative precision of 140.590% and a heritability of 0.456 (Table 4.10). The average plot height in the experiment was 635.519 mm, where the top five candidates were 38 (734.997 mm), 45 (726.172 mm), 54 (724.231 mm), 46 (723.757 mm) and 35 (710.519) (Table 4.17). The top five candidates were significantly higher than the average. It would be favourable to use this trait as a selection tool for future crossing parents.

Table 0.17: Top five candidates for the height of the plots

Rank	HYLD Entry number	Mean plot height in field (mm)	CV (%)
1	38	734.997	9.7
2	45	726.172	5.4
3	54	724.231	9.6
4	46	723.757	4.7
5	35	710.519	3.4

The RCBD moisture results were good with a CV-value of 1.12%, R^2-value of 0.5461 and H^2-value of 0.134 (Table 4.9). The NNA results had a CV-value of 0.895%, relative precision of 111.102% and H^2-value of 0.370 (Table 4.10). The NNA were a better option as it indicated that field trends do have an influence on this trait. The average of the trait was 11.514% where the top five candidates were 18 (11.705%), 15 (11.696%), 56 (11.586%), 31 (11.684%) and 41 (11.682%) (Table 4.18). The moisture of these candidates was not significantly higher than the average. It would be ineffective to select for this trait as most of the values are similar and they are not significant from the average.

Table 0.18: Top five candidates for the moisture

Rank	HYLD Entry number	Mean moisture (%)	CV (%)
1	18	11.705	0.4
2	15	11.696	0.3
3	56	11.685	1.1
4	31	11.684	1.1
5	41	11.682	0.7

The plant height trait results for the RBCD had a CV-value of 9.16%, R^2-value of 0.4821 and H^2-value of 0.212 (Table 4.9). The NNA results had CV value of 8.157%, relative precision of 122.957% and H^2-value of 0.297 (Table 4.10). The NNA indicated that field trend had an influence on this trait and should be used. The average height of the plants was 598.497 mm, where the top five candidates were 45 (714.234 mm), 10 (679.697 mm), 46 (669.456 mm), 35 (662.992 mm) and 56 (661.441 mm) (Table 4.19). The best candidate were entries 45, 10 and 46 since they were the top three candidates and also had the lowest CV-values of 2.8%, 1.7% and 2.4% respectfully.

Table 0.19: Top five candidates for the height of the individual plants

Rank	HYLD Entry number	Mean plant height (mm)	CV (%)
1	45	714.234	2.8
2	10	679.697	1.7
3	46	669.456	2.4
4	35	662.992	5.5
5	56	661.441	8.5

The RCBD for the protein trait had good results as it had a CV-value of 11.14%, R^2-value of 0.5940 and a H^2-value of 0.289 (Table 4.9). The NNA had a CV-value of 7.013%, relative precision if 192.223% and H^2-value of 0.439 (Table 4.10). The NNA was the better analysis as it indicated that field trends did have an influence on this trait. The average of the experiment for this trait was 12.238%, where the top five candidates were 3 (14.076%), 42 (14.046%) 60 (13.895%), 39 (13.797%) and 1 (13.670%). It would be beneficial to select for this trait as it has a low CV-value (Table 4.20). The top five candidates were significantly higher than the average and the best candidate was 39 with the very low CV-value of 0.9%.

Table 0.20: Top five candidates for protein

Rank	HYLD Entry number	Mean protein (%)	CV (%)
1	3	14.076	7.5
2	42	14.046	5.4
3	60	13.895	5.9
4	39	13.797	0.9
5	1	13.670	6.8

The RCBD results for dry protein had a CV-value of 11.19%, R^2-value of 0.5936 and a H^2-value of 0.289 (Table 4.9). The NNA had a CV-value of 7.032, relative precision of 192.860% and an H^2-value of 0.441 (Table 4.10). The NNA were a better option and indicated that there were field trends that influenced the trait. The average of the entries was 13.904%, where the top five candidates were 3 (15.983%), 42 (15.903%), 60 (15.840%), 39 (15.738%) and 1 (15.515%) (Table 4.20). The best candidate for this trait was entry 39 since it had a very low CV-value of 0.9%.

Table 0.21: Top five candidates for the dry protein

Rank	HYLD Entry number	Mean protein dry (%)	CV (%)
1	3	15.983	7.6
2	42	15.903	5.5
3	60	15.840	5.6
4	39	15.738	0.9
5	1	15.515	6.8

The RCBD results for the seeds per tiller trait had a CV-value of 23.46%, R^2-value of 0.4868 and H^2-value of 0.192 (Table 4.9). The NNA had a CV-value of 22.079, relative precision of 110.172% and an H^2-value of 0.2304 (Table 4.10). The NNA indicated field trends and were the better analysis for the trait. The average amount of seeds was 24.674 were the top five candidates were 48 (44.153), 13 (32.871), 54 (31.981), 55 (31.585) and 7 (31.091) (Table 4.22). The top five entries had significantly more seeds per tiller than the average. It would not be effective to select for this trait as it had high CV-values.

Table 0.22: Top five candidates for the total seeds per tiller

Rank	HYLD Entry number	Mean seeds	CV (%)
1	48	44.153	24.3
2	13	32.871	17.0
3	54	31.981	11.2
4	55	31.585	34.4
5	7	31.091	27.0

The RCBD results for the seed area had a CV-value of 8.53%, R^2-value of 0.5009 and H^2-value of 0.111 (Table 4.9). The NNA had CV-value of 7.929%, relative precision of 113.048% and a H^2-value of 0.183 (Table 4.10). The NNA indicated that field trends did have an influence on this trait. The average seed area was 7.385 mm^2, where the best five candidates were 35 (8.657 mm^2), 59 (8.375 mm^2), 26 (8.356 mm^2), 36 (8.353 mm^2) and 46 (8.143 mm^2) (Table 4.23). The best five candidates were significantly higher than the average. This is a good trait to select for since the CV values are low. The best candidates were entry 59 and 26 since they had the lowest CV-value of 3.2% and 4.9% respectfully.

Table 0.23: Top five candidates for the total seed area

Rank	HYLD Entry number	Mean seed area (mm^2)	CV (%)
1	35	8.657	6.4
2	59	8.375	3.2
3	26	8.356	4.9
4	36	8.353	5.8
5	46	8.143	6.0

The RCBD results for the seed length trait had good results with a CV-value of 3.26%, R^2-value of 0.6747 and H^2-value of 0.397 (Table 4.9). The NNA had a CV-value of 2.991%, relative precision of 113.763% and a H^2-value of 0.460 (Table 4.10). The average length of the seeds were 4.847 mm, where the top five candidates were 40 (5.292 mm), 46 (5.234 mm), 35 (5.221 mm), 60 (5.185 mm) and 26 (5.157 mm) (Table 4.24). The seeds of the top five candidates were significantly longer than the average. This trait is good to select for in breeding programmes.

Table 0.24: Top five candidates for the seed length

Rank	HYLD Entry number	Mean seed length (mm)	CV (%)
1	40	5.292	3.9
2	46	5.234	3.6
3	35	5.221	3.9
4	60	5.185	3.0
5	26	5.157	3.3

The RCBD for the seed weight trait had a CV-value of 30.80%, a R^2-value of 0.6747 and H^2-value of 0.092 (Table 4.9). The NNA had a CV-value of 26.199%, relative precision of 126.355 and H^2-value of 0.254 (Table 4.10). The NNA indicated that there were field trends that influenced the trait. The average seed weight per tiller was 0.504

g, where the top five candidates were 34 (0.717 g), 36 (0.708 g), 59 (0.696 g), 33 (0.670 g) and 7 (0.658 g) (Table 4.25). The weight of the top five candidates were significantly higher than the average. It would not be beneficial to select for this trait in the experiment since the CV-values were very high.

Table 0.25: Top five candidates for the total seed weight

Rank	HYLD Entry number	Mean seed weight (g)	CV (%)
1	34	0.717	15.2
2	36	0.708	22.4
3	59	0.696	31.5
4	33	0.670	24.2
5	7	0.658	14.6

The results of the RCBD for the seed width trait had a CV-value of 5.23%, R^2-value of 0.4956 and H^2-value of 0.160 (Table 4.9). The NNA for this trait had a CV-value of 4.447%, relative precision of 124.288% and an H^2-value of 0.386 (Table 4.10). The NNA indicated that field trends influenced this trait and this analysis should be used. The average seed width was 1.961 mm with the top five candidates were entry 35 (2.162 mm), 55 (2.144 mm), 34 (2.113 mm), 48 (2.091 mm) and 19 (2.080 mm) (Table 4.26). The best candidate to select for this trait was entry 48 since it had a CV value of 0.6% and this trait can be used as a selection tool for increased yield.

Table 0.26: Top five candidates for the total seed width

Rank	HYLD Entry number	Mean seed width (mm)	CV (%)
1	35	2.162	4.3
2	55	2.144	2.1
3	34	2.113	4.3
4	48	2.091	0.6
5	19	2.080	3.5

The RCBD results for the total spikelets trait displayed a CV-value of 15.765%, R^2-value of 0.4265 and an H^2-value of 0.127 (Table 4.9). The NNA for this trait had a CV-value of 14.412%, relative precision of 115.235% and a H^2-value of 0.211 (Table 4.10). The NNA was the better option as field trends did have an influence on the data. The average amount of spikelets per tiller was 11.440 with the top five candidates being entry 48 (14.201), 15 (14.048), 13 (13.963). 32 (13.692) and 2 (13.372) (Table 4.27).

The best two candidates to select would be 15 and 32 since they have CV-values of 7.6% and 9.3% respectfully.

Table 0.27: Top five candidates for the total spikelets

Rank	HYLD Entry number	Mean total spikelets	CV (%)
1	48	14.201	10.6
2	15	14.048	7.6
3	13	13.963	14.9
4	32	13.692	9.3
5	2	13.372	23.3

The RCBD for the spike length had a CV-value of 14.07%, R^2-value of 0.5900 and an H^2-value of 0.251 (Table 4.9). The NNA results displayed a CV-value of 12.398%, relative precision of 128.247 and a H^2-value of 0.359 (Table 4.10). The NNA was the better option as it indicated that the trait was influenced by field trends. The average spike length was 61.242 mm with the top five candidates being entry 15 (73.094 mm), 26 (72.741 mm), 60 (72.265 mm), 48 (72.193 mm) and 49 (71.483 mm) (Table 4.28). This trait is a good selection tool to increase the yield and entry 15 and 26 would be the best to select for since they were the two best candidates and had the lowest CV-values of 6.9% and 5.3%.

Table 0.28: Top five candidates for the spike length trait

Rank	HYLD Entry number	Mean spike length (mm)	CV (%)
1	15	73.094	6.9
2	26	72.741	5.3
3	60	72.265	12.9
4	48	72.193	11.9
5	49	71.483	14.1

The RCBD for the TKW had a CV-value of 18.57%, R^2-value of 0.4413 and an H^2-value of 0.114 (Table 4.9). The NNA had a CV-value of 15.791%, relative precision of 126.552% and an H^2-value of 0.252 (Table 4.10). The average TKW was 18.688 g with the top five candidates being 38 (23.248 g), 58 (22.934 g), 59 (22.694 g), 12 (22.668 g) and 46 (22.254 g) (Table 4.29). All of the top five candidates had significantly higher values than the average. The overall CV-value was very high for this trait and it would not be recommended to use this as a selection tool.

Table 0.29: Top five candidates for the TKW trait

Rank	HYLD Entry number	Mean TKW (g)	CV (%)
1	38	23.248	20.6
2	58	22.934	21.1
3	59	22.694	21.4
4	12	22.668	15.8
5	46	22.254	6.1

The RCBD for the total tillers per plant had a CV-value of 25.73%, R^2-value of 0.4347 and H^2-valuie of 0.115 (Table 4.9). The NNA had a CV-value of 24.340%, relative precision of 101.336% and an H^2-value of 0.201 (Table 4.10). The average amount of tillers per plant was 3.032 with the top five candidates with the highest number of tillers being 12 (4.243), 3 (4.111), 25 (4.045), 18 (3.910) and 17 (3.778) (Table 4.30). This trait would not be beneficial to select for this experiment as the top five candidates were not significantly higher than the average and the CV-values are very high.

Table 0.30: Top five candidates for the total tillers per plant

Rank	HYLD Entry number	Mean total tiller per plant	CV (%)
1	12	4.243	1.8
2	3	4.111	9.4
3	25	4.045	24.8
4	18	3.910	2.0
5	17	3.778	5.1

The RCBD results for the wet gluten trait had a CV-value of 12.99%, R^2-value of 0.5930 and H^2-value of 0.288 (Table 4.9). The NNA results for this trait had a CV-value of 27.828%, relative precision of 193.445% and H^2-value of 0.437 (Table 4.10). The NNA analysis indicated that there were field trends and this analysis was the better option to correct for it. The average for the wet gluten trait was 27.828% with the top five candidates being entry 3 (32.682%), 42 (32.565%), 60 (32.221%), 39 (32.067%) and 1 (31.626%) (Table 4.31). The top five candidates had significantly higher wet gluten amount than the average. The best candidate to select would be entry 39 as it was in the top five and has a low CV-value of 0.9%.

Table 0.31: Top five candidates for the wet gluten

Rank	HYLD Entry number	Mean wet gluten (%)	CV (%)
1	3	32.682	8.6
2	42	32.565	6.1
3	60	32.221	6.6
4	39	32.067	0.9
5	1	31.626	7.7

Chapter 5: Conclusion

Higher yield in wheat is an important factor for farmers and breeders, because of the increase in the human population. The aim of this study was to identify traits related to higher yield in wheat and to use them in the MS-MARS facilitated pre-breeding program. Potential donor germplasm with high yield-related traits was obtained and the lines with the best yield-related traits were selected for molecular characterization. The selected lines were screened for rust resistance, semi-dwarfing, baking quality and yield-related gene markers. The molecular markers along with the molecular diversity assessment could indicate if a desired gene was present and data were used to select the donors.

MS-MARS cycles were conducted in each season during this study where the male fertile population was used to fertilize the male sterile population. The first cycle generated noticeably more seed than the second which resulted from the male fertile population being planted too late to pollinate the male sterile population. The success rate in percentage of the cross-pollination for the first and second cycle was 23.95% and 14.44%. This can be improved by selecting the tillers at the correct reproductive time and arranging the tillers better. The structures of the florets could also have prevented pollen to efficiently cross-pollinate.

The recurrent base population of the MS-MARS scheme has an expected 1:1 ratio male sterile:male fertile. Statistical analysis of the first cycle indicated that the single dominant gene within a heterozygous situation that controls male sterility within the population could not be confirmed. In the second cycle the single dominant gene could be confirmed and the ratio was as expected. The divergence from the ratio in the first cycle is possibly due to human error.

Phenotypic data were collected when the field trial was at anthesis. During this experiment hand-held equipment was used to measure the LAI, stomatal conductance, chlorophyll content and leaf temperature of each entry in the field trial. At the same day an RPAS was used to capture images using a Parrot Sequoia multispectral camera and a GoPro with a flat RGB lens and a NI lens to calculate NDVI. Orthomosaic images were generated from the raw images and used to calculate spectral indices. R-squared

values of the linear regression models indicated that the data that was captured by the RPAS could not predict the phenotypic data in the field, except for the yield. Scatter plot matrices also showed that there were no correlation between the RPAS data and the individual ground measurements, also except for the yield.

When the plants in the field trial was ready for harvest, three single plants was pulled from each entry to determine if they possess high yielding traits. Each tiller of the plants were measured individually by hand, machines, scales, rulers and a mobile application. A RCBD and NNA were done on the data to identify which of the entries did the best in each high-yielding trait. The NNA were the better option as it indicated that the field trends had influences on most of the high-yielding traits.

In this study wheat with high yielding traits were identified and introduced into the MS-MARS at SU-PBL. Detailed observations were made by using HTPPs in the field as well as post-harvest to select the best lines for the scheme. Future work should entail adding more molecular markers that correlates with yield-related traits. The development of cameras and software used to do the phenotypical analysis of the yield-related traits will also increase the speed and precision of the collection of data. Different remote sensing equipment and software such as thermal cameras can be used to provide information on plant heights and soil moisture. Additional vegetative indices can also be explored in future studies that can be used for various reasons such as to better distinguish wheat from different plants. The model can be improved over time, with the addition of data, to develop a pipeline that can really benefit the industry. Overall there are more challenges in this field, such as flight time of the RPAS and the resolution of the satellites doing remote sensing, but research and development can have a huge contribution.

Chapter 6: References

3DR. (2020). 3DR - Secure drone and data solutions for construction, mining, and government | 3DR. [online] Available at: https://3dr.com/ [Accessed 3 Jan. 2020].

Abdullah, A., Aziz, M., Siddique, K. and Flower, K. (2015). Film antitranspirants increase yield in drought stressed wheat plants by maintaining high grain number. Agricultural water management, 159, pp.11-18.

Abhinandan, K., Skori, L., Stanic, M., Hickerson, N., Jamshed, M. and Samuel, M. (2018). Abiotic stress signaling in wheat – An inclusive overview of hormonal interactions during abiotic stress responses in wheat. Frontiers in plant science, 9.

Acevedo, E., Silva, P. and Silva, H. (2002). Wheat growth and physiology. In Curtis, B.C., Rajaram, S. and Gómez, H. Macpherson [eds.], Bread Wheat: Improvement and Production. FAO Plant Production and Protection Series No. 30.

Acquaah, G., (2015). Conventional plant breeding principles and techniques. In: Advances in plant breeding strategies: breeding, biotechnology and molecular tools. Springer international publishing, Cham, pp.115–158.

Aisawi, K.A.B., Reynolds, M.P., Singh, R.P., and Foulkes M.J. (2015). The physiological basis of the genetic progress in yield potential of CIMMYT spring wheat cultivars from 1966 to 2009. Crop Science, 55: pp.1749-1764.

Ali, M. (2011). Pedigree selection for grain yield in spring wheat (*Triticum aestivum L.*) under drought stress conditions. Asian journal of crop science, 3(4), pp.158-168.

Allen, A., Winfield, M., Burridge, A., Downie, R., Benbow, H., Barker, G., Wilkinson, P., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., Pirani, A., Webster, T., Brew, F., Bloor, C., Griffiths, S., Bentley, A., Alda, M., Jack, P., Phillips, A. and Edwards, K. (2016). Characterization of a wheat breeders' array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). Plant biotechnology journal. 15(3), pp.390-401.

Anon, (2020). [online] Available at: https://www.thingiverse.com/thing:1674587 [Accessed 3 Jan. 2020].

Araus, J. and Cairns, J. (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends in plant science, 19(1), pp.52-61.

Atkinson, N. and Urwin, P. (2012). The interaction of plant biotic and abiotic stresses: from genes to the field. Journal of Experimental Botany, 63(10), pp.3523-3543.

Atwell, W.A., (2001). An overview of wheat development, cultivation and production. Cereal Foods World, 46, 2, pp.59-62.

Baresel, J., Rischbeck, P., Hu, Y., Kipp, S., Hu, Y., Barmeier, G., Mistele, B. and Schmidhalter, U. (2017). Use of a digital camera as alternative method for non-destructive detection of the leaf chlorophyll content and the nitrogen nutrition status in wheat. Computers and electronics in agriculture, 140, pp.25-33.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M. and Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International journal of applied earth observation and geoinformation, 39, pp.79-87.

Blum, A., Shpiler, L., Golan, G. and Mayer, J. (1989). Yield stability and canopy temperature of wheat genotypes under drought-stress. Field crops research, 22(4), pp.289-296.

Braun, H.J., Atlin, G. and Payne, T., (2010). Multi-location testing as a tool to identify plant response to global climate change. Climate change and crop production, pp.115-138.

Breseghello, F. and Coelho, A. (2013). Traditional and modern plant breeding methods with examples in rice (*Oryza sativa L.*). Journal of Agricultural and Food Chemistry, 61(35), pp.8277-8286.

Bruno, E., Choi, Y., Chung, I. and Kim, K. (2017). QTLs and analysis of the candidate gene for amylose, protein, and moisture content in rice (*Oryzasativa L.*). 3 Biotech, 7(1).

Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., Maurer, H., Reif, J., Würschum, T., Müller, J., Rahe, F. and Ruckelshausen, A. (2013).

BreedVision — A multi-sensor platform for non-destructive field-based phenotyping in plant breeding. Sensors, 13(3), pp.2830-2847.

Castagliola, P., Achouri, A., Taleb, H., Celano, G. and Psarakis, S. (2012). Monitoring the coefficient of variation using a variable sampling interval control chart. Quality and reliability engineering international, 29(8), pp.1135-1149.

CGIAR. (2020). CGIAR: Science for humanity's greatest challenges. [online] Available at: https://www.cgiar.org/ [Accessed 2 Jan. 2020].

Chauhan, H., Boni, R., Bucher, R., Kuhn, B., Buchmann, G., Sucher, J., Selter, L., Hensel, G., Kumlehn, J., Bigler, L., Glauser, G., Wicker, T., Krattinger, S. and Keller, B. (2015). The wheat resistance gene Lr34 results in the constitutive induction of multiple defense pathways in transgenic barley. The plant journal, 84(1), pp.202-215.

Chhetri, M., Bariana, H., Wong, D., Sohail, Y., Hayden, M. and Bansal, U. (2017). Development of robust molecular markers for marker-assisted selection of leaf rust resistance gene Lr23 in common and durum wheat breeding programs. Molecular breeding, 37(3).

Christopher, J., Christopher, M., Jennings, R., Jones, S., Fletcher, S., Borrell, A., Manschadi, A., Jordan, D., Mace, E. and Hammer, G. (2013). QTL for root angle and number in a population developed from bread wheat (*Triticum aestivum*) with contrasting adaptation to water-limited environments. Theoretical and Applied Genetics, 126(6), pp.1563-1574.

Cimmyt.org. (2018). CIMMYT. International Maize and Wheat Improvement Center | CIMMYT. International Maize and Wheat Improvement Center. [online] Available at: http://www.cimmyt.org/ [Accessed 21 Sep. 2018].

Cossani, C.M. and Reynolds, M.P., (2012). Physiological traits for improving heat tolerance in wheat. Plant physiology, 160(4), pp.1710-1718.

Dai, J., Bean, B., Brown, B., Bruening, W., Edwards, J., Flowers, M., Karow, R., Lee, C., Morgan, G., Ottman, M., Ransom, J. and Wiersma, J. (2016). Harvest index and straw yield of five classes of wheat. Biomass and bioenergy, 85, pp.223-227.

Del Pozo, A., Yáñez, A., Matus, I., Tapia, G., Castillo, D., Sanchez-Jardón, L. and Araus, J. (2016). Physiological traits associated with wheat yield potential and performance under water-stress in a mediterranean environment. Frontiers in plant science, 7.

Desjardins, A. and Hohn, T. (1997). Mycotoxins in plant pathogenesis. Molecular plant-Microbe Interactions, 10(2), pp.147-152.

Desta, Z. and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. Trends in Plant Science, 19(9), pp.592-601.

De Vries, A. (1971). Flowering biology of wheat, particularly in view of hybrid seed production — A review. Euphytica, 20(2), pp.152-170.

Dreisigacker, S. (2012). Genetic marker systems in wheat breeding. In physiological breeding I: interdisciplinary approaches to improve crop adaptation, pp.60-68.

Duveiller, E., Singh, P.K., Mezzalama, M., Singh, R.P. and Dababat, A. (2012). Wheat diseases and pests: A guide for field identification, second edition. Mexico: CIMMYT.

Echávarri, B. and Cistué, L. (2015). Enhancement in androgenesis efficiency in barley (Hordeumvulgare L.) and bread wheat (*Triticum aestivum L.)* by the addition of dimethyl sulfoxide to the mannitol pretreatment medium. Plant cell, tissue and organ culture (PCTOC), 125(1), pp.11-22.

Ellis, J., Lagudah, E., Spielmeyer, W. and Dodds, P. (2014). The past, present and future of breeding rust resistant wheat. Frontiers in plant science, 5.

Ellis, M., Spielmeyer, W., Gale, K., Rebetzke, G. and Richards, R., (2002). " Perfect" markers for the *Rht-B1b* and *Rht-D1b* dwarfing genes in wheat. TAG Theoretical and Applied Genetics, 105(6), pp.1038-1042.

Fan, X., Kawamura, K., Xuan, T., Yuba, N., Lim, J., Yoshitoshi, R., Minh, T., Kurokawa, Y. and Obitsu, T. (2017). Low-cost visible and near infrared camera on an unmanned aerial vehicle for assessing the herbage biomass and leaf area index in an Italian ryegrass field. Grassland science. 64(2), pp.145-150.

Fischer, R.A., (2011). Wheat physiology: a review of recent developments. Crop and pasture science, 62(2), pp.95-114.

Foulkes, M.J., Reynolds, M.P. and Sylvester-Bradley, R., (2009). Genetic improvement of grain crops: yield potential. Crop physiology: applications for genetic improvement and agronomy, pp.355-385.

Frenkel, O., Jaiswal, A., Elad, Y., Lew, B., Kammann, C. and Graber, E. (2017). The effect of biochar on plant diseases: what should we learn while designing biochar substrates. Journal of environmental engineering and landscape management, 25(2), pp.105-113.

Fuentes, S., Poblete-Echeverría, C., Ortega-Farias, S., Tyerman, S. and De Bei, R. (2014). Automated estimation of leaf area index from grapevine canopies using cover photography, video and computational analysis methods. Australian journal of grape and wine research, 20(3), pp.465-473.

Gandhi, G., Parthiban, S., Thummalu, N. and Christy, A. (2015). NDVI: vegetation change detection using remote sensing and gis – a case study of vellore district. Procedia computer science, 57, pp.1199-1210.

Gao, L., Wang, A., Li, X., Dong, K., Wang, K., Appels, R., Ma, W. and Yan, Y. (2009). Wheat quality related differential expressions of albumins and globulins revealed by two-dimensional difference gel electrophoresis (2-D DIGE). Journal of Proteomics, 73(2), pp.279-296.

Garcia, G.A., Hasan, A.K., Puhl, L.E., Reynolds, M.P., Calderini, D.F. and Miralles, D.J., (2013). Grain yield potential strategies in an elite wheat double-haploid population grown in contrasting environments. Crop Science, 53(6), pp.2577-2587.

Gegas, V., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., Sayers, L., Doonan, J. and Snape, J. (2010). A genetic framework for grain size and shape variation in wheat. The plant cell, 22(4), pp.1046-1056.

Gonzalez-Dugo, V., Hernandez, P., Solis, I. and Zarco-Tejada, P. (2015). Using high-resolution hyperspectral and thermal airborne imagery to assess physiological

condition in the context of wheat phenotyping. Remote sensing, 7(10), pp.13586-13605.

Gorjanc, G., Jenko, J., Hearne, S. and Hickey, J. (2016). Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. BMC genomics, 17(1).

Groos, C., Robert, N., Bervas, E. and Charmet, G. (2003). Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. Theoretical and applied genetics, 106(6), pp.1032-1040.

Guo, Z. and Schnurbusch, T., (2015). Variation of floret fertility in hexaploid wheat revealed by tiller removal. Journal of experimental botany, 66(19), pp.5945-5958.

Guo, Z., Song, Y., Zhou, R., Ren, Z. and Jia, J. (2009). Discovery, evaluation and distribution of haplotypes of the wheat *Ppd-D1* gene. New phytologist, 185(3), pp.841-851.

Guzman, C., Peña, R., Singh, R., Autrique, E., Dreisigacker, S., Crossa, J., Rutkoski, J., Poland, J. and Battenfield, S. (2016). Wheat quality improvement at CIMMYT and the use of genomic selection on it. Applied and translational genomics, 11, pp.3-8.

Haghighattalab, A., González Pérez, L., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., Ortiz-Monasterio, I., Singh, R., Goodin, D. and Poland, J. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. Plant methods, 12(1).

Havé, M., Leitao, L., Bagard, M., Castell, J. and Repellin, A. (2015). Protein carbonylation during natural leaf senescence in winter wheat, as probed by fluorescein-5-thiosemicarbazide. Plant biology, 17(5), pp.973-979.

He, J., Zhao, X., Laroche, A., Lu, Z., Liu, H. and Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Frontiers in plant science, 5.

Helguera, M., Khan, I.A., Kolmer, J., Lijavetzky, D., Zhong-qi, L. & Dubcovsky, J. (2003). PCR assays for the *Lr37-Yr17-Sr38* cluster of rust resistance genes and

their use to develop isogenic hard red spring wheat lines. Crop Science, 43 pp.1839-1847

Hemdane, S., Jacobs, P., Dornez, E., Verspreet, J., Delcour, J. and Courtin, C. (2015). Wheat (*Triticum aestivum L.*) Bran in Bread Making: A critical review. Comprehensive reviews in food science and food safety, 15(1), pp.28-42.

Herndl, M., White, J., Hunt, L., Graeff, S. and Claupein, W. (2008). Erratum to "Field-based evaluation of vernalisation requirement, photoperiod response and earliness per se in bread wheat (*Triticum aestivum L.*)" [Field crop res. 105 (2008) 193–201]. Field crops research, 107(1), p.87.

Hörtensteiner, S. (2009). Chlorophyll retention regulates chlorophyll and chlorophyll-binding protein degradation during senescence. Trends in plant science, 14(3), pp.155-162.

Hospital, F. (2009). Challenges for effective marker-assisted selection in plants. Genetica, 136(2), pp.303-310.

Hsu, W., Davis, J., Chipman, R. and Pau, S. (2015). Compound dichroic polarizers with wavelength-dependent transmission axes. Applied optics, 54(21), p.6476.

Hunag, M.-C.; Wu, J.; Cang, J.; Yang, D. (2015). An efficient k-means clustering algorithm using simple partitioning. J. Inf. Sci. Eng., 21(6), pp.1157–1177.

Iqbal, M., Shahzad, A. and Ahmed, I., (2011). Allelic variation at the *Vrn-A1*, *Vrn-B1*, *Vrn-D1*, *Vrn-B3* and *Ppd-D1a* loci of Pakistani spring wheat cultivars. Electronic journal of biotechnology, 14(1), pp.1-2.

Iwyp.org. (2020). International Wheat Yield Partnership » International Wheat Yield Partnership. [online] Available at: https://iwyp.org/ [Accessed 2 Jan. 2020].

Jackson, R., Idso, S., Reginato, R. and Pinter, P. (1981). Canopy temperature as a crop water stress indicator. Water resources research, 17(4), pp.1133-1138.

Keeble-Gagnère, G., Rigault, P., Tibbits, J., Pasam, R., Hayden, M., Forrest, K., Frenkel, Z., Korol, A., Huang, B., Cavanagh, C., Taylor, J., Abrouk, M., Sharpe, A., Konkin, D., Sourdille, P., Darrier, B., Choulet, F., Bernard, A., Rochfort, S.,

Dimech, A., Watson-Haigh, N., Baumann, U., Eckermann, P., Fleury, D., Juhasz, A., Boisvert, S., Nolin, M., Doležel, J., Šimková, H., Toegelová, H., Šafář, J., Luo, M., Câmara, F., Pfeifer, M., Isdale, D., Nyström-Persson, J., IWGSC, Koo, D., Tinning, M., Cui, D., Ru, Z. and Appels, R. (2018). Optical and physical mapping with local finishing enables megabase-scale resolution of agronomically important regions in the wheat genome. Genome Biology, 19(1).

Keller, B., Lagudah, E.S., Selter,L.L., Risk, J.M., Harsh, C and Krattinger, S.G. (2013). How has Lr34/Yr18 conferred effective rust resistance in wheat for so long? [Online] Available: http://www.globalrust.org/sites/default/files/BGRI-2012-plenary. [Accessed 21 Feb. 2020].

Khush, G. (2001). Green revolution: the way forward. Nature reviews genetics, 2(10), pp.815-822.

Kim, H., Kim, S., Kim, J., Jin, K. and Kim, H. (2015). Generating selected color using RGB, auxiliary lights, and simplex search. MATEC web of conferences, 32, p.05007.

Kipp, S., Mistele, B. and Schmidhalter, U. (2014). Identification of stay-green and early senescence phenotypes in high-yielding winter wheat, and their relationship to grain yield and grain protein concentration using high-throughput phenotyping techniques. Functional plant biology, 41(3), p.227.

Komyshev, E., Genaev, M. and Afonnikov, D. (2018). Evaluation of the SeedCounter, a mobile application for grain phenotyping. Frontiers in Plant Science, 7.

Krattinger, S.G., Lagudah, E.S., Spielmeyer, W., Singh, R.P., Huerta-Espino, J., McFadden, H., Bossolini, E., Selter, L.L. and Keller, B. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science, 323, pp.1360-1363.

Kumar, S., Singh, S.S., Mishra, C.N., Saroha, M., Gupta, V., Sharma, P., Tiwari, V. and Sharma, I., (2015). Assessment of tiller inhibition (*tin*) gene molecular marker for its application in marker-assisted breeding in wheat. National academy science letters, 38(6), pp.457-460.

Lagudah, E. (2011). Molecular genetics of race non-specific rust resistance in wheat. Euphytica, 179(1), pp.81-91.

Lagudah, E.S., McFadden, H., Singh, R.P., Huerta-Espino, J., Bariana, H.S. and Spielmeyer, W., (2006). Molecular genetic characterization of the *Lr34/Yr18* slow rusting resistance gene region in wheat. Theoretical and Applied Genetics, 114(1), pp.21-30.

Langridge, P., (2014). Reinventing the green revolution by harnessing crop mutant resources. Plant physiology, 166(4), pp.1682-1683.

Lodish H, Berk A, Zipursky SL, Molecular Cell Biology. 4th edition. New York: W. H. Freeman; (2000). Section 16.3, Photosynthetic stages and light-absorbing pigments. Available from: https://www.ncbi.nlm.nih.gov/books/NBK21598/ [Accessed 2 Feb. 2020].

Longnecker, N., Kirby, E. and Robson, A. (1993). Leaf emergence, tiller growth, and apical development of nitrogen-dificient spring wheat. Crop science, 33(1), p.154.

Lopes, M. and Reynolds, M. (2012). Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. Journal of experimental botany, 63(10), pp.3789-3798.

Luo, Y., Guan, K. and Peng, J. (2018). STAIR: A generic and fully-automated method to fuse multiple sources of optical satellite data to generate a high-resolution, daily and cloud-/gap-free surface reflectance product. Remote sensing of environment, 214, pp.87-99.

Ma, L., Li, T., Hao, C., Wang, Y., Chen, X. and Zhang, X. (2015). *TaGS5-3A*, a grain size gene selected during wheat improvement for larger kernel and yield. Plant biotechnology journal, 14(5), pp.1269-1280.

Madec, S., Baret, F., de Solan, B., Thomas, S., Dutartre, D., Jezequel, S., Hemmerlé, M., Colombeau, G. and Comar, A. (2017). High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground lidar estimates. Frontiers in plant science, 8.

Mago, R., Bariana, H.S., Dundas, I.S., Spielmeyer, W., Lawrence, G.J., Pryor, A.J. and Ellis, J.G. (2005). Development of PCR markers for the selection of wheat stem rust resistance genes *Sr24* and *Sr26* in diverse wheat germplasm. Theoretical and Applied Genetics, 111, pp.496-504.

Mago, R., Brown-Guedira, G., Dreisigacker, S., Breen, J., Jin, Y., Singh, R., Appels, R., Lagudah, E.S., Ellis, J. and Spielmeyer, W., (2011). An accurate DNA marker assay for stem rust resistance gene *Sr2* in wheat. Theoretical and applied genetics, 122(4), pp.735- 744.

Marais, G.F. and Botes, W.C. (2009). Recurrent mass selection for routine improvement of common wheat. In: E. Lichtfouse (ed). Organic farming, pest control and remediation of soil pollutants. Sustainable agricultural reviews. Springer science and business media, pp 85-105.

Marais, G.F., Botes, W.C. and Louw, J.H., (2000). Recurrent selection using male sterility and hydroponic tiller culture in pedigree breeding of wheat. Plant breeding, 119(5), pp.440-442.

Marsalis M.A., Goldberg N.P., (2016). Leaf, stem, and stripe rust diseases of wheat. New Mexico State University. Guide A: 1-6

Massa, G., Wheeler, R., Morrow, R. and Levine, H. (2016). Growth chambers on the International Space Station for large plants. Acta horticulturae, (1134), pp.215-222.

Matsuoka, Y. (2011). Evolution of polyploid *Triticum* wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant and cell physiology, 52(5), pp.750–764.

Montazeaud, G., Karatoğma, H., Özturk, I., Roumet, P., Ecarnot, M., Crossa, J., Özer, E., Özdemir, F. and Lopes, M. (2016). Predicting wheat maturity and stay–green parameters by modeling spectral reflectance measurements and their contribution to grain yield under rainfed conditions. Field crops research, 196, pp.191-198.

Montesinos-López, O., Montesinos-López, A., Crossa, J., de los Campos, G., Alvarado, G., Suchismita, M., Rutkoski, J., González-Pérez, L. and Burgueño, J.

(2017). Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. Plant methods, 13(1).

Nardini, A., Ounapuu-Pikas, E. and Savi, T. (2014). When smaller is better: leaf hydraulic conductance and drought vulnerability correlate to leaf size and venation density across four *Coffeaarabica* genotypes. Functional plant biology, 41(9), p.972.

Niroula, R. K., and Bimb, H. P. (2009). Overview of wheat X maize system of crosses for dihaploid induction in wheat. World applied sciences journal, 7(8), pp.1037-1045.

Niu, Z., Jiang, A., Abu Hammad, W., Oladzadabbasabadi, A., Xu, S. S., Mergoum, M., and Elias, E. M. (2014). Review of doubled haploid production in durum and common wheat through wheatÅ~ maize hybridization. Plant breeding, 133(3), pp.313-320.

Ortiz-Monasterio, J.I., Sayre, K.D., Rajaram, S. and McMahon, M., (1997). Genetic progress in wheat yield and nitrogen use efficiency under four nitrogen rates. Crop sci. 37, pp.898-904.

Otsu, N. A threshold selection method from gray-level histograms. IEEE Trans. Syst. man cybern. 1979, 9, pp.62–66.

Oyiga, B., Sharma, R., Baum, M., Ogbonnaya, F., Léon, J. and Ballvora, A. (2017). Allelic variations and differential expressions detected at quantitative trait loci for salt stress tolerance in wheat. Plant, cell and environment. 41(5), pp.919-935.

Piepho, H. and Möhring, J. (2007). Computing heritability and selection response from unbalanced plant breeding trials. Genetics, 177(3), pp.1881-1888.

Pignone, D., De Paola, D., Rapanà, N. and Janni, M. (2015). Single seed descent: a tool to exploit durum wheat (*Triticum durum Desf*.) genetic resources. Genetic resources and crop evolution, 62(7), pp.1029-1035.

Prins, R., Groenewald, J.Z., Marais, G.F., Snape, J.W. and Koebner, R.M.D., (2001) (AFLP and STS tagging of *Lr19*, a gene conferring resistance to leaf rust in wheat. Theoretical and Applied Genetics, 103(4), pp.618-624.

Qin, L., Hao, C., Hou, J., Wang, Y., Li, T., Wang, L., Ma, Z. and Zhang, X. (2014). Homologous haplotypes, expression, genetic effects and geographic distribution of the wheat yield gene *TaGW2*. BMC plant biology, 14(1), p.107.

Rahman, M., Chikushi, J., Yoshida, S. and Karim, A. (2009). Growth and yield components of wheat genotypes exposed to high temperature stress under control environment. Bangladesh journal of agricultural research, 34(3).

Rana, B., Rana, P., Yadav, M.K. and Kumar, S. (2011). Marker assisted selection strategy for wheat improvement. Department of Biotechnology, Sardar Vallabh Bhai Patel University of Agriculture and Technology, Meerut U.P., India, pp.19-30.

Raymond Hunt, E. and Daughtry, C. (2014). Chlorophyll meter calibrations for chlorophyll content using measured and simulated leaf transmittances. Agronomy journal, 106(3), p.931.

Reif, J., Zhang, P., Dreisigacker, S., Warburton, M., van Ginkel, M., Hoisington, D., Bohn, M. and Melchinger, A. (2005). Wheat genetic diversity trends during domestication and breeding. Theoretical and applied genetics, 110(5), pp.859-864.

Reynolds, M., Manes, Y., and Rebetzke, G. (2012). Application of physiology in breeding for heat and drought stress. In Physiological breeding I: interdisciplinary approaches to improve crop adaptation, pp. 18-32.

Rhoda, R. (2018). Improving wheat grain yield by employing an integrated biotechnology approach. MSc. Stellenbosch University.

Richardson, A., Duigan, S. and Berlyn, G. (2002). An evaluation of noninvasive methods to estimate foliar chlorophyll content. New Phytologist, 153(1), pp.185-194.

Röder, M., Korzun, V., Gill, B. and Ganal, M. (1998). The physical mapping of microsatellite markers in wheat. Genome, 41(2), pp.278-283.

Rout, P., Naik, N., Ngangkham, U., Verma, R., Katara, J., Singh, O. and Samantaray, S. (2016). Doubled Haploids generated through anther culture from an elite long

duration rice hybrid, CRHR32: Method optimization and molecular characterization. Plant biotechnology, 33(3), pp.177-186.

SAGIS, South African grain information service.http://www.sagis.org.za/, accessed 26 Dec. 2019.

Sharma, S.N., Sain, R.S. and Sharma, R.K., (2003). Genetics of spike length in durum wheat. Euphytica, 130(2), pp.155-161.

Siddique, K., Kirby, E. and Perry, M. (1989). Ear: Stem ratio in old and modern wheat varieties; relationship with improvement in number of grains per ear and yield. Field crops research, 21(1), pp.59-78.

Simmonds, J., Scott, P., Brinton, J., Mestre, T.C., Bush, M., Del Blanco, A., Dubcovsky, J. and Uauy, C., (2016). A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. Theoretical and Applied Genetics, 129(6), pp.1099-1112.

Simmons, S.R., Oelke, E.A. and Anderson, P.M. (1995). Growth and development guide for spring wheat. University of Minnesota. [Online] Available: http://www.extension.umn.edu/distribution/cropsystems/dc2547.html Accessed: 15 May 2018.

Sinclair, T. and Jamieson, P. (2006). Grain number, wheat yield, and bottling beer: An analysis. Field crops research, 98(1), pp.60-67.

Ślusarkiewicz-Jarzina, A., Pudelska, H., Woźna, J. and Pniewski, T. (2017). Improved production of doubled haploids of winter and spring triticale hybrids via combination of colchicine treatments on anthers and regenerated plants. Journal of applied genetics, 58(3), pp.287-295.

Spielmeyer, W. and Richards, R.A., (2004). Comparative mapping of wheat chromosome 1AS which contains the tiller inhibition gene (*tin*) with rice chromosome 5S. Theoretical and applied genetics, 109(6), pp.1303-1310.

Spindel, J., Begum, H., Akdemir, D., Virk, P., Collard, B., Redoña, E., Atlin, G., Jannink, J. and McCouch, S. (2015). Correction: Genomic selection and association mapping in rice (*Oryzasativa*): Effect of trait genetic architecture, training

population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLOS genetics, 11(6), p.e1005350.

Sucher, J., Boni, R., Yang, P., Rogowsky, P., Büchner, H., Kastner, C., Kumlehn, J., Krattinger, S. and Keller, B. (2016). The durable wheat disease resistance gene *Lr34* confers common rust and northern corn leaf blight resistance in maize. Plant biotechnology journal, 15(4), pp.489-496.

Sun, J., Rutkoski, J., Poland, J., Crossa, J., Jannink, J. and Sorrells, M. (2017). Multitrait, random regression, or simple repeatability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield. The plant genome, 10(2).

Taylor, C. (2019). How do we determine outliers in statistics?. [online] ThoughtCo. Available at: https://www.thoughtco.com/what-is-an-outlier-3126227 [Accessed 5 Mar. 2019].

Tee, T. and Qualset, C. (1975). Bulk populations in wheat breeding: comparison of single-seed descent and random bulk methods. Euphytica, 24(2), pp.393-405.

Trethowan, R.M., Reynolds, M.P., Ortiz-monasterio, J.I. and Ortiz, R., 2007. The genetic basis of the green revolution in wheat production. Plant breed. Rev. 28, pp.39-58.

USDA.gov. (2019). USDA. [online] Available at: https://www.usda.gov/ [Accessed 12 Dec. 2019].

Van Zyl, J. (2017). Canopy temperature as a water stress indicator in vines. South African journal of ecology and viticulture, 7(2).

Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P. and Hawkesford, M. (2017). Field scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. Functional plant biology, 44(1), p.143.

Wan, Y., Poole, R., Huttly, A., Toscano-Underwood, C., Feeney, K., Welham, S., Gooding, M., Mills, C., Edwards, K., Shewry, P. and Mitchell, R. (2008).

Transcriptome analysis of grain development in hexaploid wheat. BMC Genomics, 9(1), p.121.

Wang, J., van Ginkel, M., Podlich, D., Ye, G., Trethowan, R., Pfeiffer, W., DeLacy, I., Cooper, M. and Rajaram, S. (2003). Comparison of two breeding strategies by computer simulation. Crop science, 43(5), p.1764.

Whingwiri, E. and Kemp, D. (1980). Spikelet development and grain yield of the wheat ear in response to applied nitrogen. Australian journal of agricultural research, 31(4), p.637.

Würschum, T., Langer, S., Longin, C., Tucker, M. and Leiser, W. (2017). A modern green revolution gene for reduced height in wheat. The plant journal, 92(5), pp.892-903.

Würschum, T., Liu, G., Boeven, P., Longin, C., Mirdita, V., Kazman, E., Zhao, Y. and Reif, J. (2018). Exploiting the *Rht* portfolio for hybrid wheat breeding. Theoretical and applied genetics.

Xu, Y. and Crouch, J.H., 2008. Marker-assisted selection in plant breeding: from publications to practice. Crop science, 48(2), pp.391-407.

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J.L., Echenique, V. and Dubcovsky, J., (2004). The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science, 303(5664), pp.1640-1644.

Zhang, Y., Liu, J., Xia, X. and He, Z. (2014). *TaGS-D1*, an ortholog of rice *OsGS3*, is associated with grain weight and grain length in common wheat. Molecular breeding, 34(3), pp.1097-1107.

Zikhali, M. and Griffiths, S., (2015). The effect of Earliness per se (*Eps*) genes on flowering time in bread wheat. In Advances in Wheat Genetics: From genome to field, pp.339-345.